A Short Guide to
Celestial Navigation

Revised April 10th, 2006

First Published May 20th, 1997

Legal Notice

1. License and Ownership

The publication “A Short Guide to Celestial Navigation“ is owned and copyrighted by

I grant the user a free nonexclusive license to download said publication for personal or educational use as well as for any lawful non-commercial purposes provided the terms of this agreement are met. To deploy the publication in any commercial context, either in-house or externally, you must obtain a written permission from the author.

2. Distribution

You may distribute copies of the publication to third parties provided you do not do so for profit and each copy is unaltered and complete. Extracting graphics and/or parts of the text is prohibited.

3. Warranty Disclaimer

I provide the publication to you "as is" without any express or implied warranties of any kind including, but not limited to, any implied warranties of fitness for a particular purpose. I do not warrant that the publication will meet your requirements or that it is error-free. You assume full responsibility for the selection, possession, and use of the publication and for verifying the results obtained by application of formulas, procedures, diagrams, or other information given in said publication.

January 1, 2006

Henning Umland

Felix qui potuit boni
fontem visere lucidum,
felix qui potuit gravis
terrae solvere vincula.

Boethius

Preface

Why should anybody still practice celestial navigation in the era of electronics and GPS? One might as well ask why some photographers still develop black-and-white photos in their darkroom instead of using a high-color digital camera. The answer would be the same: because it is a noble art, and because it is rewarding. No doubt, a GPS navigator is a powerful tool, but using it becomes routine very soon. In contrast, celestial navigation is an intellectual challenge. Finding your geographic position by means of astronomical observations requires knowledge, judgement, and skillfulness. In other words, you have to use your brains. Everyone who ever reduced a sight knows the thrill I am talking about. The way is the goal.

It took centuries and generations of navigators, astronomers, geographers, mathematicians, and instrument makers to develop the art and science of celestial navigation to its present level, and the knowledge thus acquired is a treasure that should be preserved. Moreover, celestial navigation gives an impression of scientific thinking and creativeness in the pre-electronic age. Last but not least, celestial navigation may be a highly appreciated alternative if a GPS receiver happens to fail.

When I read my first book on navigation many years ago, the chapter on celestial navigation with its fascinating diagrams and formulas immediately caught my particular interest although I was a little deterred by its complexity at first. As I became more advanced, I realized that celestial navigation is not nearly as difficult as it seems to be at first glance. Studying the literature, I found that many books, although packed with information, are more confusing than enlightening, probably because most of them have been written by experts and for experts. Other publications are designed like cookbooks, i. e., they contain step-by-step instructions but avoid much of the theory. In my opinion, one can not *really* comprehend celestial navigation and enjoy the beauty of it without knowing the mathematical background.

Since nothing really complied with my needs, I decided to write a compact manual for my personal use which had to include the most important definitions, formulas, diagrams, and procedures. The idea to publish it came in 1997 when I became interested in the internet and found that it is the ideal medium to share one's knowledge with others. I took my manuscript, rewrote it in the HTML format, and published it on my web site. Later, I converted everything to the PDF format, which is an established standard for electronic publishing now.

The style of my work may differ from standard books on this subject. This is probably due to my different perspective. When I started the project, I was a newcomer to the world of navigation, but I had a background in natural sciences and in scientific writing. From the very beginning, it has been my goal to provide accurate information in a highly structured and comprehensible form. The reader may judge whether this attempt has been successful.

More people than I ever expected are interested in celestial navigation, and I would like to thank readers from all over the world for their encouraging comments and suggestions. However, due to the increasing volume of correspondence, I am no longer able to answer every individual question or to provide individual support. Unfortunately, I have still a few other things to do, e. g., working for my living. Nonetheless, I keep working on this publication at leisure.

I apologize for misspellings, grammar errors, and wrong punctuation. I did my best, but English is not my native language.

Last but not least, I owe my wife an apology for spending countless hours in front of the PC, staying up late, neglecting household chores, etc. I'll try to mend my ways. Some day ...

April 10th, 2006

Henning Umland

Correspondence address:

Dr. Henning Umland
Rabenhorst 6
21244 Buchholz i. d. N.
Germany

Fax: +49 721 151296535

Index

Chapter 1

The Basics of Celestial Navigation

Celestial navigation, a branch of applied astronomy, is the art and science of finding one's geographic position through astronomical observations, particularly by measuring **altitudes** of celestial bodies – sun, moon, planets, or stars.

An observer watching the night sky without knowing anything about geography and astronomy might spontaneously get the impression of being on a plane located at the center of a huge, hollow sphere with the celestial bodies attached to its inner surface. Indeed, this naive model of the universe was in use for millennia and developed to a high degree of perfection by ancient astronomers. Still today, it is a useful tool for celestial navigation since the navigator, like the astronomers of old, measures **apparent positions** of bodies in the sky but not their absolute positions in space.

The apparent position of a body in the sky is defined by the **horizon system of coordinates**. In this system, the observer is located at the center of a fictitious hollow sphere of infinite diameter, the **celestial sphere**, which is divided into two hemispheres by the plane of the **celestial horizon** (*Fig. 1-1*). The **altitude**, **H**, is the vertical angle between the line of sight to the respective body and the celestial horizon, measured from 0° through +90° when the body is above the horizon (visible) and from 0° through -90° when the body is below the horizon (invisible). The **zenith distance**, **z**, is the corresponding angular distance between the body and the **zenith**, an imaginary point vertically overhead. The zenith distance is measured from 0° through 180°. The point opposite to the zenith is called **nadir** (z = 180°). H and z are **complementary angles** (H + z = 90°). The **true azimuth**, **Az_N**, is the horizontal direction of the body with respect to the geographic (true) north point on the horizon, measured clockwise from 0° through 360°.

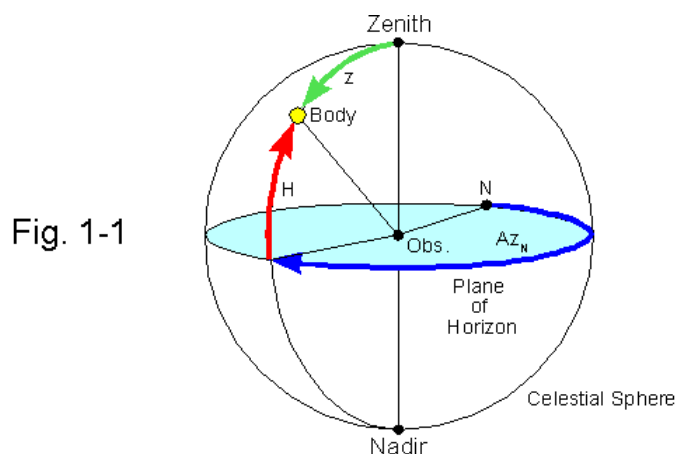
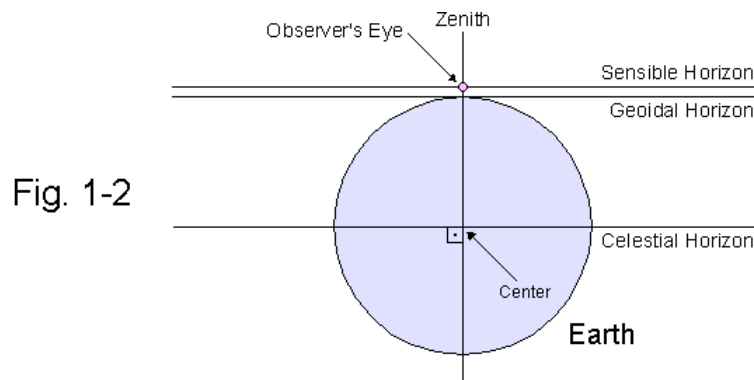
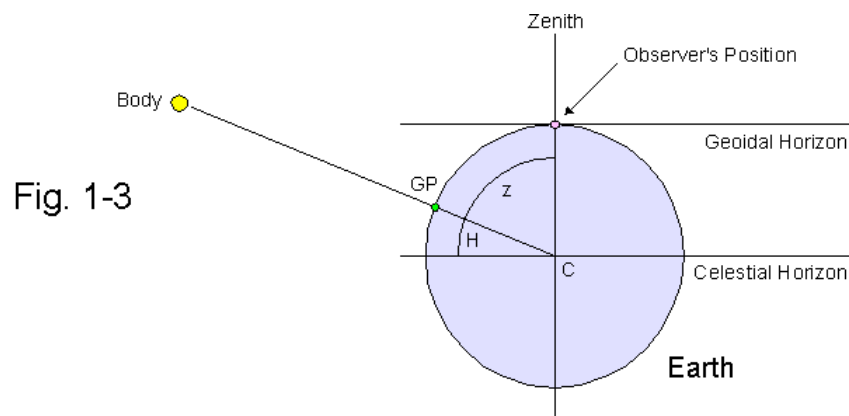
Fig. 1-1

In reality, the observer is **not** located on the plane of the celestial horizon. *Fig. 1-2* shows the three horizontal planes relevant to celestial navigation.

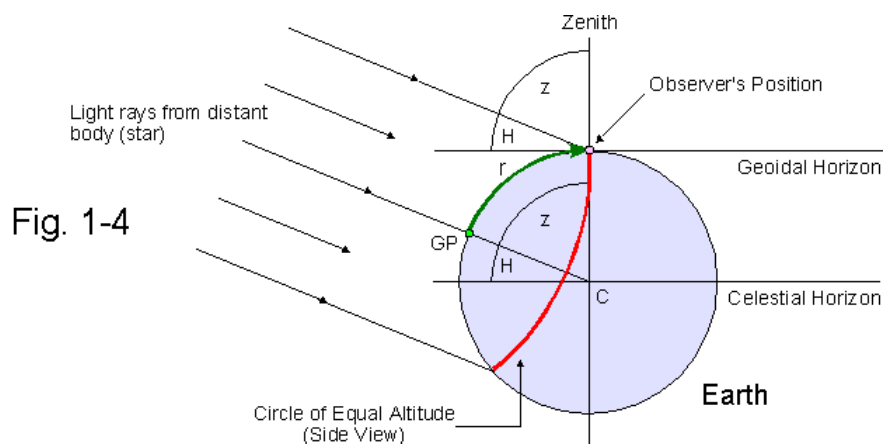
Fig. 1-2

The **celestial horizon** is the horizontal plane passing through the center of the earth which coincides with the center of the celestial sphere. The **geoidal horizon** is the horizontal plane tangent to the earth at the observer's position. The **sensible horizon** is the horizontal plane passing through the observer's eye. These three planes are parallel to each other.

The sensible horizon merges into the geoidal horizon when the observer's eye is at sea level. Since both horizons are usually very close to each other, they can be considered as identical under practical conditions. None of the above horizontal planes coincides with the **visible horizon**, the line where the earth's surface and the sky appear to meet.
Calculations of celestial navigation always refer to the geocentric altitude of a body, the altitude measured by a fictitious observer being on the plane of the celestial horizon and at the center of the earth which coincides with the center of the celestial sphere.

Since there is no way to measure the geocentric altitude directly, it has to be derived from the altitude with respect to the visible or sensible horizon (altitude corrections, chapter 2).

A **marine sextant** is an instrument designed to measure the altitude of a body with reference to the visible sea horizon. Instruments with any kind of an **artificial horizon** measure the altitude referring to the sensible horizon (chapter 2).

Altitude and zenith distance of a celestial body depend on the distance between the terrestrial observer and the **geographic position of the body**, **GP**. GP is the point where a straight line from the celestial body to the center of the earth, C, intersects the earth's surface (*Fig. 1-3*).

A body appears in the zenith (z = 0°, H = 90°) when GP is identical with the observer's position. A terrestrial observer moving away from GP will observe that the altitude of the body decreases as his distance from GP increases. The body is on the celestial horizon (H = 0°, z = 90°) when the observer is one quarter of the circumference of the earth away from GP.

For a given altitude of a body, there is an infinite number of positions having the same distance from GP and forming a circle on the earth's surface whose center is on the line C–GP (*Fig 1-4*), below the earth's surface. Such a circle is called a **circle of equal altitude**. An observer traveling along a circle of equal altitude will measure a constant altitude and zenith distance of the respective body, no matter where on the circle he is. The radius of the circle, r, measured along the surface of the earth, is directly proportional to the observed zenith distance, z.

$$r\left[nm\right] = 60 \cdot z\left[°\right] \qquad or \qquad r\left[km\right] = \frac{Perimeter\ of\ Earth\left[km\right]}{360°} \cdot z\left[°\right]$$

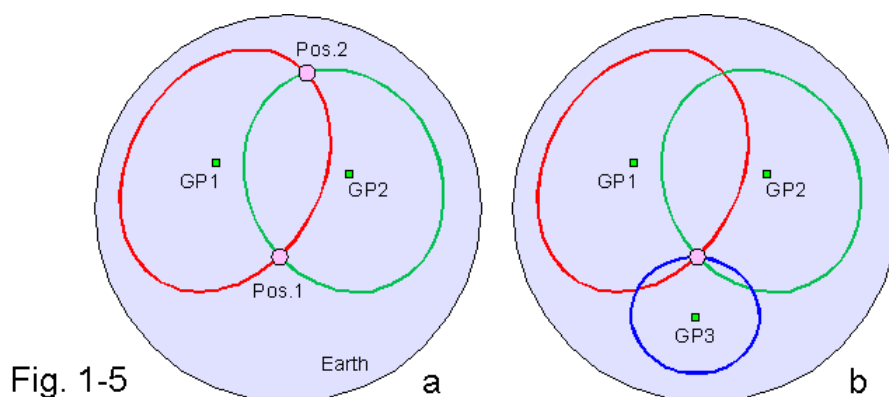
One nautical mile (1 nm = 1.852 km) is the **great circle distance** of one minute of arc (the definition of a great circle is given in chapter 3). The mean perimeter of the earth is 40031.6 km.

Light rays originating from distant objects (stars) are virtually parallel to each other when reaching the earth. Therefore, the altitude with respect to the geoidal (sensible) horizon equals the altitude with respect to the celestial horizon. In contrast, light rays coming from the relatively close bodies of the solar system are diverging. This results in a measurable difference between both altitudes (parallax). The effect is greatest when observing the moon, the body closest to the earth (see chapter 2, *Fig. 2-4*).

The azimuth of a body depends on the observer's position on the circle of equal altitude and can assume any value between 0° and 360°.

Whenever we measure the altitude or zenith distance of a celestial body, we have already gained partial information about our own geographic position because we know we are somewhere on a circle of equal altitude with the radius r and the center GP, the geographic position of the body. Of course, the information available so far is still incomplete because we could be anywhere on the circle of equal altitude which comprises an infinite number of possible positions and is therefore also called a **circle of position** (see chapter 4).

We continue our thought experiment and observe a second body in addition to the first one. Logically, we are on two circles of equal altitude now. Both circles overlap, intersecting each other at two points on the earth's surface, and one of those two points of intersection is our own position (*Fig. 1-5a*). Theoretically, both circles could be tangent to each other. For several reasons, however, this case is undesirable and has to be avoided (see chapter 16).

Fig. 1-5

In principle, it is not possible to know which point of intersection – Pos.1 or Pos.2 – is identical with our actual position unless we have additional information, e.g., a fair estimate of where we are, or the **compass bearing** of at least one of the bodies. Solving the problem of **ambiguity** can also be achieved by observation of a third body because there is only one point where all three circles of equal altitude intersect (*Fig. 1-5b*).

Theoretically, we could find our position by plotting the circles of equal altitude on a globe. Indeed, this method has been used in the past but turned out to be impractical because precise measurements require a very big globe. Plotting circles of equal altitude on a map is possible if their radii are small enough. This usually requires observed altitudes of almost 90°. The method is rarely used since such altitudes are not easy to measure. In most cases, circles of equal altitude have diameters of several thousand nautical miles and can **not** be plotted on usual maps. Further, plotting circles on a map is made more difficult by geometric distortions related to the respective map projection (chapter 13).

Since a navigator always has an estimate of his position, it is not necessary to plot whole circles of equal altitude but rather those parts which are near the expected position.

In the 19th century, navigators developed graphic methods for the construction of lines (secants and tangents of the circles of equal altitude) whose point of intersection marks the observer's position. These revolutionary methods, which are considered as the beginning of modern celestial navigation, will be explained later.

In summary, finding one's position by astronomical observations includes three basic steps:

1. Measuring the altitudes or zenith distances of two or more chosen bodies (chapter 2).

2. Finding the geographic position of each body at the time of its observation (chapter 3).

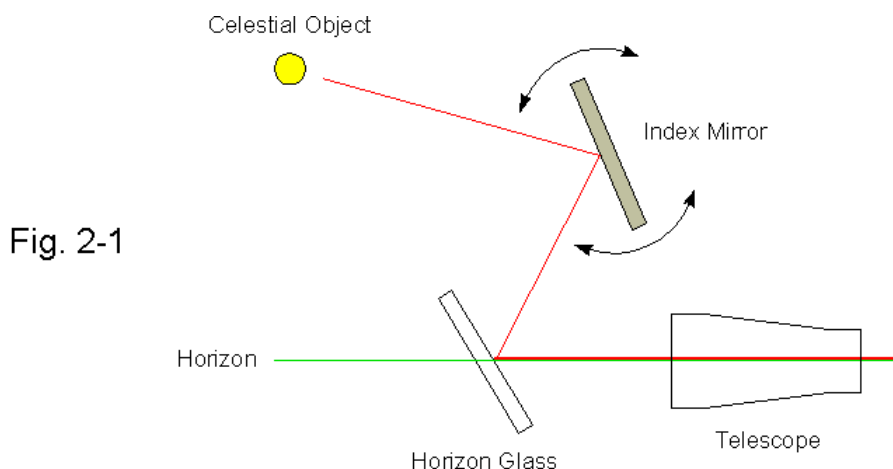
3. Deriving one's own position from the above data (chapter 4&5).

Chapter 2

Altitude Measurement

Although altitudes and zenith distances are equally suitable for navigational calculations, most formulas are traditionally based upon altitudes since these are easily accessible using the visible sea horizon as a natural reference line. Direct measurement of the zenith distance, however, requires an instrument with an artificial horizon, e.g., a pendulum or spirit level indicating the direction of gravity (perpendicular to the local plane of horizon), since a reference point in the sky does not exist.

Instruments

A **marine sextant** consists of a system of two mirrors and a telescope mounted on a metal frame. A schematic illustration (side view) is given in *Fig. 2-1*. The rigid horizon glass is a semi-translucent mirror attached to the frame. The fully reflecting index mirror is mounted on the so-called index arm rotatable on a pivot perpendicular to the frame. When measuring an altitude, the instrument frame is held in a vertical position, and the visible sea horizon is viewed through the scope and horizon glass. A light ray coming from the observed body is first reflected by the index mirror and then by the back surface of the horizon glass before entering the telescope. By slowly rotating the index mirror on the pivot the superimposed image of the body is aligned with the image of the horizon. The corresponding altitude, which is twice the angle formed by the planes of horizon glass and index mirror, can be read from the graduated limb, the lower, arc-shaped part of the sextant frame (not shown). Detailed information on design, usage, and maintenance of sextants is given in [3] (see appendix).

On land, where the horizon is too irregular to be used as a reference line, altitudes have to be measured by means of instruments with an artificial horizon:

A **bubble attachment** is a special sextant telescope containing an **internal artificial horizon** in the form of a small spirit level whose image, replacing the visible horizon, is superimposed with the image of the body. Bubble attachments are expensive (almost the price of a sextant) and not very accurate because they require the sextant to be held absolutely still during an observation, which is difficult to manage. A sextant equipped with a bubble attachment is referred to as a **bubble sextant**. Special bubble sextants were used for air navigation before electronic navigation systems became standard equipment.

A pan filled with water, or preferably a more viscous liquid, e. g., glycerol, can be utilized as an **external artificial horizon**. Due to the gravitational force, the surface of the liquid forms a perfectly horizontal mirror unless distorted by vibrations or wind. The vertical angular distance between a body and its mirror image, measured with a marine sextant, is twice the altitude. This very accurate method is the perfect choice for exercising celestial navigation in a backyard.

A **theodolite** is basically a telescopic sight which can be rotated about a vertical and a horizontal axis. The angle of elevation is read from the vertical circle, the horizontal direction from the horizontal circle. Built-in spirit levels are used to align the instrument with the plane of the sensible horizon before starting the observations (artificial horizon). Theodolites are primarily used for surveying, but they are excellent navigation instruments as well. Many models can measure angles to 0.1' which cannot be achieved even with the best sextants. A theodolite is mounted on a tripod and has to stand on solid ground. Therefore, it is restricted to land navigation. Traditionally, theodolites measure zenith distances. Modern models can optionally measure altitudes.

Never view the sun through an optical instrument without inserting a proper shade glass, otherwise your eye might suffer permanent damage !

Altitude corrections

Any altitude measured with a sextant or theodolite contains errors. Altitude corrections are necessary to eliminate systematic altitude errors and to reduce the altitude measured relative to the visible or sensible horizon to the altitude with respect to the celestial horizon and the center of the earth (chapter 1). Altitude corrections do not remove random errors.

Index error (IE)

A sextant or theodolite, unless recently calibrated, usually has a constant error (**index error, IE**) which has to be subtracted from the readings before they can be used for navigational calculations. The error is positive if the displayed value is greater than the actual value and negative if the displayed value is smaller. Angle-dependent errors require alignment of the instrument or the use of an individual correction table.

$$1st\ correction: \qquad H_1 = Hs - IE$$

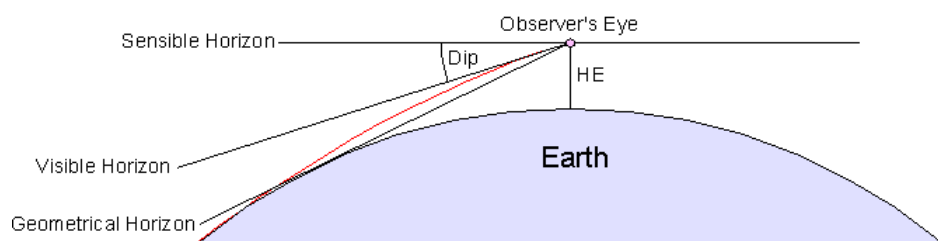
The **sextant altitude**, **Hs**, is the altitude as indicated by the sextant before any corrections have been applied.

When using an external artificial horizon, H_1 (not Hs!) has to be divided by two.

A theodolite measuring the zenith distance, z, requires the following formula to obtain H_1:

$$H_1 = 90° - (z - IE)$$

Dip of horizon

If the earth's surface were an infinite plane, visible and sensible horizon would be identical. In reality, the visible horizon appears several arcminutes below the sensible horizon which is the result of two contrary effects, the curvature of the earth's surface and atmospheric refraction. The **geometrical horizon**, a flat cone, is formed by an infinite number of straight lines tangent to the earth and radiating from the observer's eye. Since atmospheric refraction bends light rays passing along the earth's surface toward the earth, all points on the geometric horizon appear to be elevated, and thus form the visible horizon. If the earth had no atmosphere, the visible horizon would coincide with the geometrical horizon (*Fig. 2-2*).

The altitude of the sensible horizon relative to the visible horizon is called **dip** and is a function of the **height of eye**, **HE**, the vertical distance of the observer's eye from the earth's surface:

$$Dip['] \approx 1.76 \cdot \sqrt{HE[m]} \approx 0.97 \cdot \sqrt{HE[ft]}$$

The above formula is empirical and includes the effects of the curvature of the earth's surface and atmospheric refraction*.

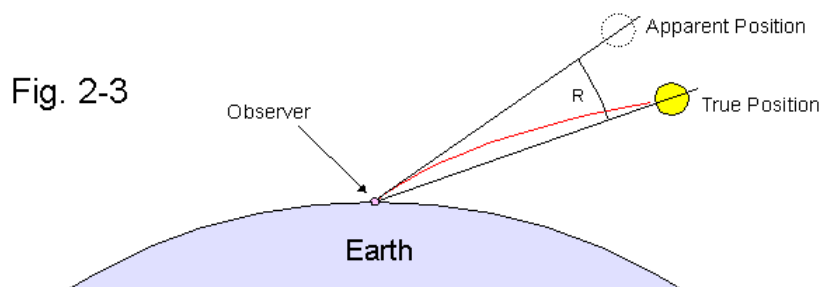
*At sea, the dip of horizon can be obtained directly by measuring the vertical angle between the visible horizon in front of the observer and the visible horizon behind the observer (through the zenith). Subtracting 180° from the angle thus measured and dividing the resulting angle by two yields the dip of horizon. This very accurate method is rarely used because it requires a special instrument (similar to a sextant).

$$2nd\ correction: \qquad H_2 = H_1 - Dip$$

The correction for dip has to be omitted (dip = 0) if any kind of an artificial horizon is used since an artificial horizon indicates the sensible horizon.

The altitude H_2 obtained after applying corrections for index error and dip is also referred to as **apparent altitude**, **Ha**.

Atmospheric refraction

A light ray coming from a celestial body is slightly deflected toward the earth when passing obliquely through the atmosphere. This phenomenon is called **refraction**, and occurs always when light enters matter of different density at an angle smaller than 90°. Since the eye can not detect the curvature of the light ray, the body appears to be at the end of a straight line tangent to the light ray at the observer's eye and thus appears to be higher in the sky. R is the angular distance between apparent and true position of the body at the observer's eye (*Fig. 2-3)*.

Refraction is a function of Ha (= H_2). Atmospheric **standard refraction**, $\mathbf{R_0}$, is 0' at 90° altitude and increases progressively to approx. 34' as the apparent altitude approaches 0°:

Ha [°]	0	1	2	5	10	20	30	40	50	60	70	80	90
$\mathbf{R_0}$ **[']**	~34	~24	~18	9.9	5.3	2.6	1.7	1.2	0.8	0.6	0.4	0.2	0.0

R_0 can be calculated with a number of formulas like, e. g., *Smart's* formula which gives highly accurate results from 15° through 90° altitude [2,9]:

$$R['] = \frac{0.97127}{\tan H_2[°]} - \frac{0.00137}{\tan^3 H_2[°]}$$

For navigation, *Smart's* formula is still accurate enough at 10° altitude. Below 5°, the error increases progressively. For altitudes between 0° and 15°, the following formula is recommended [10]. H_2 is measured in degrees:

$$R_0['] = \frac{34.133 + 4.197 \cdot H_2 + 0.00428 \cdot H_2^2}{1 + 0.505 \cdot H_2 + 0.0845 \cdot H_2^2}$$

A low-precision refraction formula including the whole range of altitudes from 0° through 90° was found by *Bennett*:

$$R_0['] \;=\; \frac{1}{\tan\left(H_2[°] + \dfrac{7.31}{H_2[°] + 4.4}\right)}$$

The accuracy is sufficient for navigational purposes. The maximum systematic error, occurring at 12° altitude, is approx. 0.07' [2]. If necessary, *Bennett's* formula can be improved (max. error: 0.015') by the following correction:

$$R_{0,\,improved}['] \;=\; R_0['] - 0.06 \cdot \sin\left(14.7 \cdot R_0['] + 13\right)$$

The argument of the sine is stated in degrees [2].

Refraction is influenced by atmospheric pressure and air temperature. The standard refraction, R_0, has to be multiplied with a correction factor, f, to obtain the refraction for a given combination of pressure and temperature if high precision is required.

$$f \;=\; \frac{p[mbar]}{1010} \cdot \frac{283}{273 + T[°C]} \;=\; \frac{p[in.Hg]}{29.83} \cdot \frac{510}{460 + T[°F]}$$

P is the atmospheric pressure and T the air temperature. **Standard conditions** (f = 1) are **1010 mbar** (29.83 in) and **10°C** (50°F). The effects of air humidity are comparatively small and can be ignored.

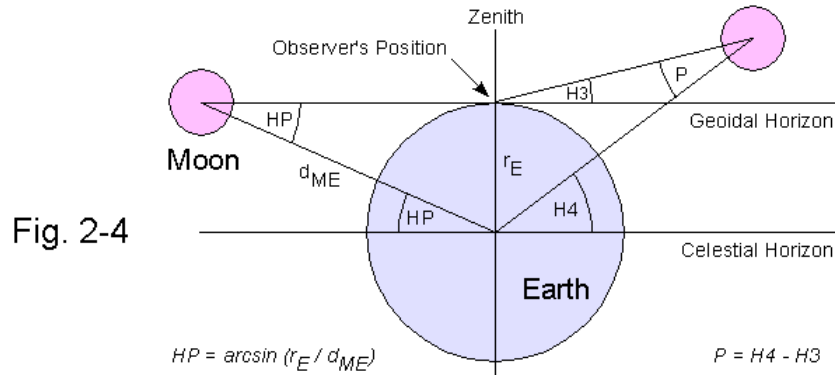
Refraction formulas refer to a fictitious standard atmosphere with the most probable density gradient. The actual refraction may differ from the calculated one if anomalous atmospheric conditions are present (temperature inversion, mirage effects, etc.). Particularly at low altitudes, anomalies of the atmosphere gain influence. Therefore, refraction at altitudes below ca. 5° may become erratic, and calculated values are not always reliable. It should be mentioned that dip, too, is influenced by atmospheric refraction and may become unpredictable under certain meteorological conditions.

$$3rd\ correction: \qquad H_3 \;=\; H_2 - f \cdot R_0$$

H_3 is the altitude of the body with respect to the sensible horizon.

Parallax

Calculations of celestial navigation refer to the altitude with respect to the earth's center and the celestial horizon. *Fig. 2-4* illustrates that the altitude of a near object, e.g., the moon, with respect to the celestial horizon, H_4, is noticeably greater than the altitude with respect to the geoidal (sensible) horizon, H_3. The difference H_4-H_3 is called **parallax in altitude**, **P**. It decreases with growing distance between object and earth and is too small to be measured when observing stars (compare with chapter 1, *Fig. 1-4*). Theoretically, the observed parallax refers to the sensible, not to the geoidal horizon.

Since the height of eye is several magnitudes smaller than the radius of the earth, the resulting error in parallax is not significant (< 0.0003' for the moon at 30 m height of eye).

The parallax (in altitude) of a body being on the geoidal horizon is called **horizontal parallax**, **HP**. The HP of the sun is approx. 0.15'. Current HP's of the moon (ca. 1°!) and the navigational planets are given in the **Nautical Almanac** [12] and similar publications, e.g., [13]. P is a function of altitude and HP of a body:

$$P = \arcsin\left(\sin HP \cdot \cos H_3\right) \approx HP \cdot \cos H_3$$

When we observe the upper or lower limb of a body (see below), we assume that the parallax of the limb equals the parallax of the center (when at the same altitude). For geometric reasons (curvature of the surface), this is not quite correct. However, even with the moon, the body with by far the greatest parallax, the resulting error is so small that it can be ignored (<< 1").

The above formula is rigorous for a spherical earth. However, the earth is not exactly a sphere but resembles an **oblate spheroid**, a sphere flattened at the poles (chapter 9)*. This may cause a small but measurable error (≤ 0.2') in the parallax of the moon, depending on the observer's position [12]. Therefore, a small correction, ΔP, should be added to P if high precision is required:

$$\Delta P \approx f \cdot HP \cdot \left[\sin\left(2 \cdot Lat\right) \cdot \cos Az_N \cdot \sin H_3 - \sin^2 Lat \cdot \cos H_3\right] \qquad f = \frac{1}{298.257}$$

$$P_{improved} = P + \Delta P$$

Lat is the observer's estimated latitude (chapter 4). $\mathbf{Az_N}$, the azimuth of the moon, is either measured with a compass (compass bearing) or calculated using the formulas given in chapter 4.

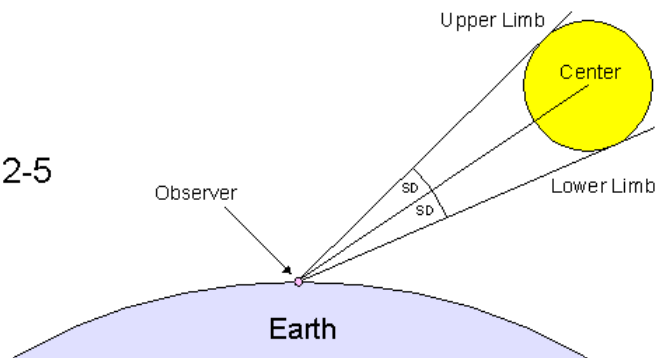
$$4th\ correction: \qquad H_4 = H_3 + P$$

* Tabulated values for HP refer to the equatorial radius of the earth (equatorial horizontal parallax).

Semidiameter

When observing sun or moon with a marine sextant or theodolite, it is not possible to locate the center of the body with sufficient accuracy. It is therefore common practice to measure the altitude of the upper or lower limb of the body and add or subtract the apparent **semidiameter**, **SD**, the angular distance of the respective limb from the center (*Fig. 2-5*).

We correct for the **geocentric** SD, the SD measured by a fictitious observer at the center the earth, since H_4 refers to the celestial horizon and the center of the earth (see *Fig. 2-4*). The geocentric semidiameters of sun and moon are given on the daily pages of the **Nautical Almanac** [12]. We can also calculate the geocentric SD of the moon from the tabulated horizontal parallax:

$$SD_{geocentric} = \arcsin\left(k \cdot \sin HP\right) \approx k \cdot HP \qquad k_{Moon} = 0.2725$$

The factor k is the ratio of the radius of the moon (1738 km) to the equatorial radius of the earth (6378 km).

Although the semidiameters of the navigational planets are not quite negligible (the SD of Venus can increase to 0.5'), the centers of these bodies are customarily observed, and no correction for SD is applied. Semidiameters of stars are much too small to be measured (SD=0).

$$5th\ correction: \qquad H_5 = H_4 \pm SD_{geocentric}$$

(lower limb: +SD, upper limb: –SD)

When using a bubble sextant which is less accurate anyway, we observe the center of the body and skip the correction for semidiameter.

The altitude obtained after applying the above corrections is called **observed altitude**, **Ho**.

$$Ho = H_5$$

Ho is the geocentric altitude of the body, the altitude with respect to the celestial horizon and the center of the earth (see chapter 1).

Alternative corrections for semidiameter and parallax

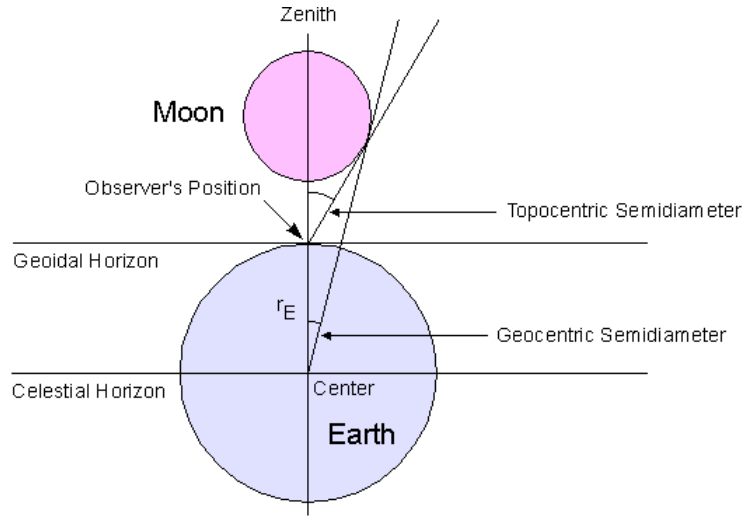
The order of altitude corrections described above is in accordance with the Nautical Almanac. Alternatively, we can correct for semidiameter **before** correcting for parallax. In this case, however, we have to calculate with the **topocentric** semidiameter, the semidiameter of the respective body as seen from the observer's position on the surface of the earth (see *Fig. 2-5*), instead of the geocentric semidiameter.

With the exception of the moon, the body nearest to the earth, there is no significant difference between topocentric and geocentric SD. The topocentric SD of the moon is only marginally greater than the geocentric SD when the moon is on the sensible horizon but increases measurably as the altitude increases because of the decreasing distance between observer and moon. The distance is smallest (decreased by about the radius of the earth) when the moon is in the zenith. As a result, the topocentric SD of the moon being in the zenith is approximately 0.3' greater than the geocentric SD. This phenomenon is called **augmentation** (Fig. 2-6).

The accurate formula for the topocentric (augmented) semidiameter of the moon is stated as:

$$SD_{topocentric} = \arctan \frac{k}{\sqrt{\frac{1}{\sin^2 HP} - \left(\cos H_3 \pm k\right)^2} - \sin H_3}$$

(observation of lower limb: +k, observation of upper limb: –k)

This formula is rigorous for a spherical earth. The error caused by the flattening of the earth is too small to be measured.

The following formula is given by *Meeus* [2]. It ignores the difference between upper and lower limb but is still accurate enough for navigational purposes (error < 1"):

$$SD_{topocentric} \approx k \cdot HP \cdot \left(1 + \sin HP \cdot \sin H_3\right)$$

A similar approximation was proposed by *Stark* [14]:

$$SD_{topocentric} \approx \frac{k \cdot HP}{1 - \sin HP \cdot \sin H_3}$$

Thus, the fourth correction is:

$$4th\ correction\ (alt.): \qquad H_{4,\,alt} = H_3 \pm SD_{topocentric}$$

(lower limb: +SD, upper limb: –SD)

$H_{4,alt}$ is the topocentric altitude of the center of the moon.

Using the parallax formulas explained above, we calculate P_{alt} from $H_{4,alt}$. Thus, the fifth correction is:

$$5th\ correction\ (alt.): \qquad H_{5,\,alt} = H_{4,\,alt} + P_{alt}$$

$$Ho = H_{5,\,alt}$$

Since the geocentric SD is easier to calculate than the topocentric SD, it is generally recommendable to correct for the semidiameter in the last place unless one has to know the augmented SD of the moon for special reasons.

Combined corrections for semidiameter and parallax of the moon

For observations of the moon, there is a surprisingly simple formula including the corrections for augmented semidiameter **as well as** parallax in altitude:

$$Ho = H_3 + \arcsin\left[\sin HP \cdot \left(\cos H_3 \pm k\right)\right]$$

(lower limb: +k, upper limb: –k)

The above formula is rigorous for a spherical earth but does not take into account the effects of the flattening. Therefore, the small correction ΔP should be added to Ho.

To complete the picture, it should be mentioned that there is also a formula to calculate the topocentric (augmented) semidiameter of the moon from the geocentric altitude of the moon's center, H:

$$SD_{topocentric} = \arcsin \frac{k}{\sqrt{1 + \frac{1}{\sin^2 HP} - 2 \cdot \frac{\sin H}{\sin HP}}}$$

This formula, too, is based upon a spherical model of the earth.

Phase correction (Venus and Mars)

Since Venus and Mars show phases similar to the moon, their apparent center may differ somewhat from the actual center. Since the coordinates of both planets tabulated in the **Nautical Almanac** [12] refer to the apparent center, an additional correction is not required. The phase correction for Jupiter and Saturn is too small to be significant.

In contrast, coordinates calculated with **Interactive Computer Ephemeris** refer to the actual center. In this case, the upper or lower limb of the respective planet should be observed if the magnification of the telescope is sufficient.

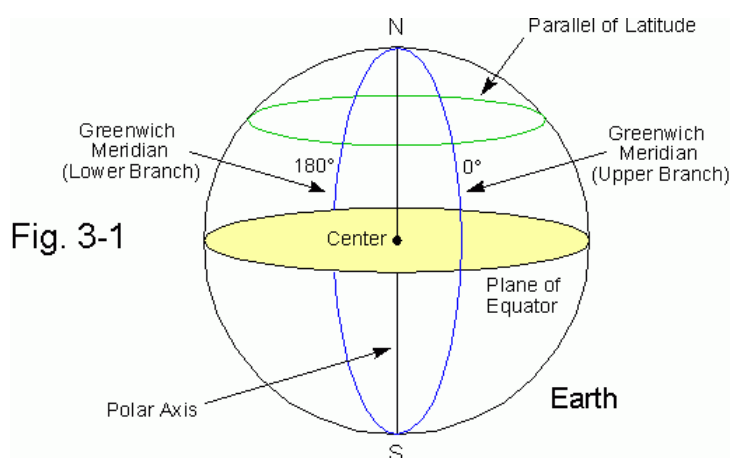
The **Nautical Almanac** provides sextant altitude correction tables for sun, planets, stars (pages A2 – A4), and the moon (pages xxxiv – xxxv), which can be used instead of the above formulas if very high precision is not required (the tables cause additional rounding errors).

Instruments with an artificial horizon can exhibit additional errors caused by acceleration forces acting on the bubble or pendulum and preventing it from aligning itself with the direction of gravity. Such acceleration forces can be random (vessel movements) or systematic (coriolis force). The coriolis force is important to air navigation and requires a special correction formula. In the vicinity of mountains, ore deposits, and other local irregularities of the earth's crust, gravity itself can be slightly deflected from the normal to the reference ellipsoid (deflection of the vertical, see chapter 9).

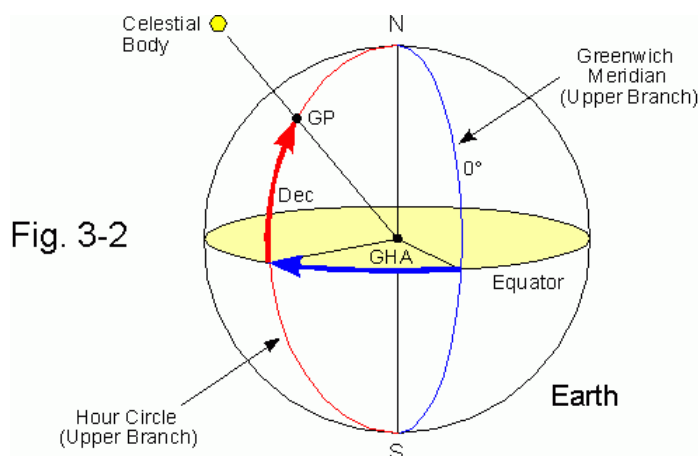
Chapter 3

Geographic Position and Time

Geographic terms

In celestial navigation, the earth is regarded as a sphere. Although this is an approximation, the geometry of the sphere is applied successfully, and the errors caused by the flattening of the earth are usually negligible (chapter 9). A circle on the surface of the earth whose plane passes through the center of the earth is called a **great circle**. Thus, a great circle has the greatest possible diameter of all circles on the surface of the earth. Any circle on the surface of the earth whose plane does **not** pass through the earth's center is called a **small circle**. The **equator** is the only great circle whose plane is perpendicular to the **polar axis**, the axis of rotation. Further, the equator is the only **parallel of latitude** being a great circle. Any other parallel of latitude is a small circle whose plane is parallel to the plane of the equator. A **meridian** is a great circle going through the **geographic poles**, the points where the polar axis intersects the earth's surface. The **upper branch** of a meridian is the half from pole to pole passing through a given point, e. g., the observer's position. The **lower branch** is the opposite half. The **Greenwich meridian**, the meridian passing through the center of the transit instrument at the **Royal Greenwich Observatory**, was adopted as the **prime meridian** at the International Meridian Conference in 1884. Its upper branch is the reference for measuring longitudes (0°...+180° east and 0°...–180° west), its lower branch (180°) is the basis for the **International Dateline** (*Fig. 3-1*).

Each point of the earth's surface has an imaginary counterpart on the surface of the celestial sphere obtained by central projection. The projected image of the observer's position, for example, is the zenith. Accordingly, there are two **celestial poles**, the c**elestial equator**, **celestial meridians**, etc.

The equatorial system of coordinates

The **geographic position of a celestial body**, **GP**, is defined by the **equatorial system of coordinates** (*Fig. 3-2*). The **Greenwich hour angle**, **GHA**, is the angular distance of GP from the upper branch of the Greenwich meridian (0°), measured **westward** from 0° through 360°. **Declination**, **Dec**, is the angular distance of GP from the plane of the equator, measured northward through +90° or southward through –90°. GHA and Dec are **geocentric coordinates** (measured at the center of the earth). The great circle going through the poles and GP is called **hour circle** (*Fig. 3-2*).

GHA and Dec are equivalent to geocentric longitude and latitude with the exception that **longitude is measured westward through −180° and eastward through +180°**.

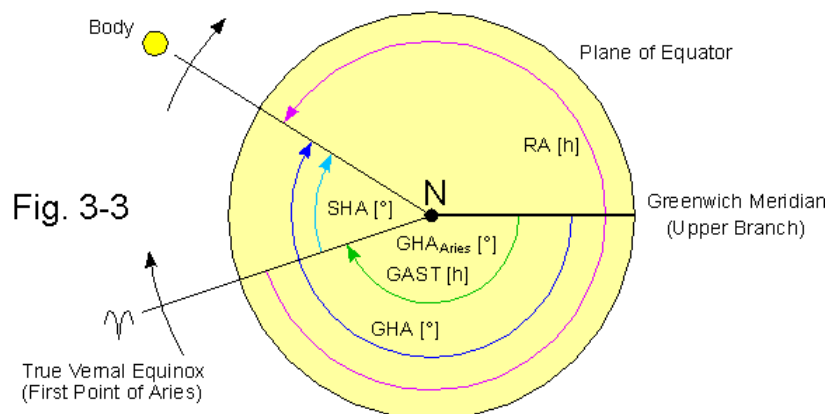
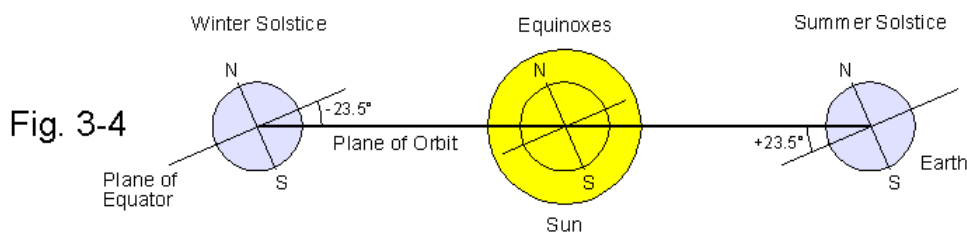
Since the Greenwich meridian rotates with the earth from west to east, whereas each hour circle remains linked with the almost stationary position of the respective body in the sky, the Greenwich hour angles of all celestial bodies increase by approx. 15° per hour (360° in 24 hours). In contrast to stars (15° 2.46' /h), the GHA's of sun, moon, and planets increase at slightly different (and variable) rates. This is caused by the revolution of the planets (including the earth) around the sun and by the revolution of the moon around the earth, resulting in additional apparent motions of these bodies in the sky. In many cases, it is useful to measure the angular distance between the hour circle of a celestial body and the hour circle of a reference point in the sky instead of the Greenwich meridian because the angle thus obtained is independent of the earth's rotation. The **sidereal hour angle**, **SHA**, is the angular distance of a body from the hour circle (upper branch) of the **first point of Aries** (also called **vernal equinox**, see below), measured **westward** from 0° through 360°. Thus, the GHA of a body is the sum of the SHA of the body and $\mathbf{GHA_{Aries}}$, the GHA of the first point of Aries, :

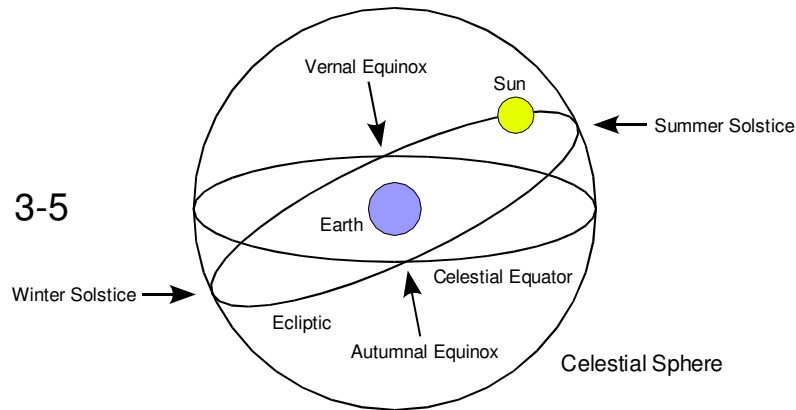
$$GHA = SHA + GHA_{Aries}$$

(If the resulting GHA is greater than 360°, subtract 360°.)

The angular distance of a body **eastward** from the hour circle of the vernal equinox, measured in time units (24h = 360°), is called **right ascension**, **RA**. Right ascension is mostly used by astronomers whereas navigators prefer SHA.

$$RA[h] = 24h - \frac{SHA[°]}{15} \quad \Leftrightarrow \quad SHA[°] = 360° - 15 \cdot RA[h]$$

Fig. 3-3 illustrates how the various hour angles are interrelated.

Declinations are not affected by the rotation of the earth. The declinations of sun and planets change primarily due to the **obliquity of the ecliptic**, the inclination of the earth's equator to the **ecliptic**. The latter is the plane of the earth's orbit and forms a great circle on the celestial sphere. The declination of the sun, for example, varies periodically between ca. +23.5° at the time of the summer solstice and ca. -23.5° at the time of the winter solstice. (*Fig.3-4*).

The two points on the celestial sphere where the great circle of the ecliptic intersects the celestial equator are called **equinoxes**. The term equinox is also used for the time at which the apparent sun, moving westward along the ecliptic, passes through either of these points, approximately on March 21 and on September 23. Accordingly, there is a **vernal equinox** (**first point of Aries, vernal point**) and an **autumnal equinox**. The former is the reference point for measuring sidereal hour angles (*Fig. 3-5*).

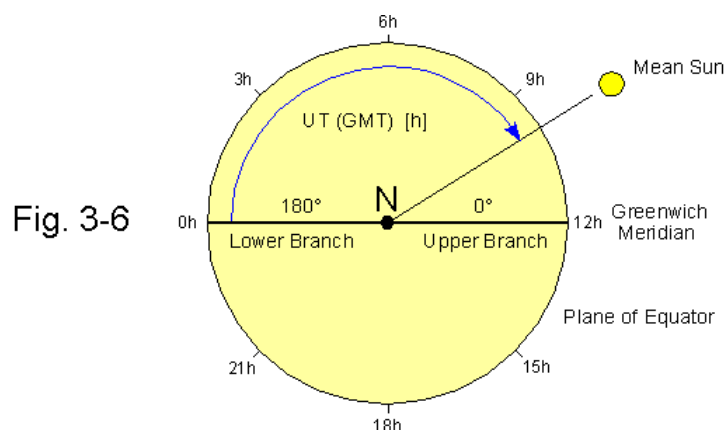
When the sun passes through either of the equinoxes (Dec ≈ 0), day and night have (approximately) the same length, regardless of the observer's position (Lat. *aequae noctes* = equal nights).

Declinations of the planets and the moon are also influenced by the inclinations of their own orbits to the ecliptic. The plane of the moon's orbit, for example, is inclined to the ecliptic by approx. 5° and makes a tumbling movement (precession, see below) with a cycle of 18.6 years (Saros cycle). As a result, the declination of the moon varies between approx. -28.5° and +28.5° at the beginning and at the end of the Saros cycle, and between approx. -18.5° and +18.5° in the middle of the Saros cycle.

Further, sidereal hour angles and declinations of all bodies change slowly due to the influence of the **precession** of the earth's polar axis. Precession is a slow, tumbling movement of the polar axis along the surface of an imaginary double cone. One revolution takes about 26000 years (Platonic year). As a result, the equinoxes move westward along the celestial equator at a rate of approx. 50" per year. Thus, the sidereal hour angle of each star decreases at about the same rate. In addition, the polar axis makes a small elliptical oscillation, called **nutation**, which causes the equinoxes to travel along the celestial equator at a periodically changing rate, and we have to distinguish between the ficticious **mean equinox of date** and the **true equinox of date** (see time measurement). Accordingly, the declination of each body oscillates. The same applies to the rate of change of the sidereal hour angle and right ascension of each body.

Even stars are not fixed in space but move individually, resulting in a slow drift of their equatorial coordinates (**proper motion**). Finally, the apparent positions of bodies are influenced by the limited speed of light (**light time, aberration**),and, to a small extent, by **annual parallax** [16]. The accurate prediction of geographic positions of celestial bodies requires complicated algorithms. Formulas for the calculation of low-precision **ephemerides** of the sun (accurate enough for celestial navigation) are given in chapter 15.

Time measurement in navigation and astronomy

Due to the rapid change of Greenwich hour angles, celestial navigation requires accurate time measurement, and the time at the instant of observation should be noted to the second if possible. This is usually done by means of a chronometer and a stopwatch. The effects of time errors are dicussed in chapter 16. On the other hand, the earth's rotation with respect to celestial bodies provides an important basis for astronomical time measurement.

Coordinates tabulated in the **Nautical Almanac** refer to **Universal Time**, **UT**. UT has replaced **Greenwich Mean Time**, **GMT**, the traditional basis for civil time keeping. Conceptually, UT (like GMT) is the hour angle of the fictitious m**ean sun**, expressed in hours, with respect to the **lower branch** of the Greenwich meridian (mean solar time, *Fig. 3-6*).

UT is calculated using the following formula:

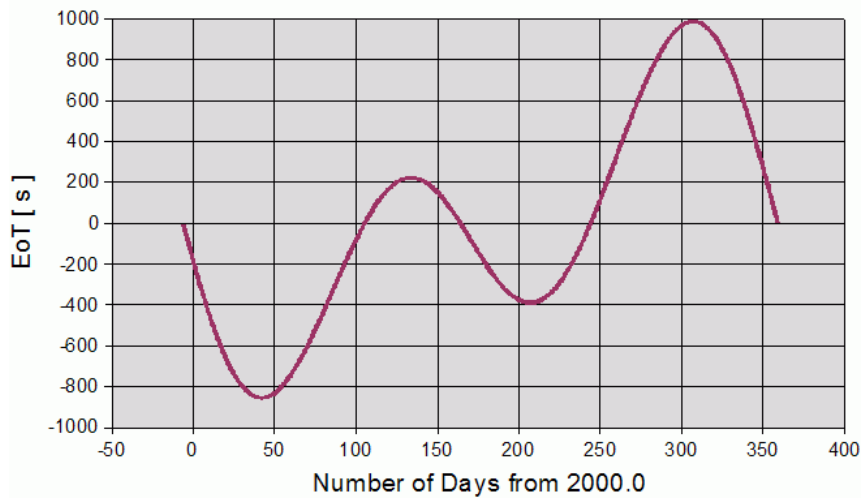
$$UT[h] = GMT[h] = \frac{GHA_{MeanSun}[°]}{15} + 12$$

(If UT is greater than 24 h, subtract 24 hours.)

By definition, the GHA of the mean sun increases by exactly 15° per hour, completing a 360° cycle in 24 hours. The unit for UT (and GMT) is 1 solar day, the time interval between two consecutive meridian transits of the mean sun.

The rate of change of the GHA of the **apparent (observable) sun** varies periodically and is sometimes slightly greater, sometimes slightly smaller than 15° per hour during the course of a year. This behavior is caused by the eccentricity of the earth's orbit and by the obliquity of the ecliptic. The time measured by the hour angle of the apparent sun with respect to the lower branch of the Greenwich meridian is called **Greenwich Apparent Time**, **GAT**. A calibrated sundial located at the Greenwich meridian, for example, would indicate GAT. The difference between GAT and UT (GMT) is called **equation of time, EoT**:

$$EoT = GAT - UT$$

EoT varies periodically between approx. −14 and +17 minutes (*Fig. 3-7*). Predicted values for EoT for each day of the year (at 0:00 and 12:00 UT) are given in the Nautical Almanac (grey background indicates negative EoT). EoT is needed when calculating times of sunrise and sunset, or determining a **noon longitude** (see chapter 6). Formulas for the calculation of EoT are given in chapter 15.

The hour angle of the mean sun with respect to the lower branch of the local meridian (the meridian going through the observer's position) is called **Local Mean Time**, **LMT**. LMT and UT are linked through the following formula:

$$LMT[h] = UT[h] + \frac{Lon[°]}{15}$$

The instant of the mean sun passing through the local meridian (upper branch) is called **Local Mean Noon**, **LMN**.

A **zone time** is the local mean time with respect to a longitude being a multiple of ±15°. Thus, zone times differ by an integer number of hours. In the US, for example, Eastern Standard Time (UT−5h) is LMT at −75° longitude, Pacific Standard Time (UT−8h) is LMT at −120° longitude. Central European Time (UT+1h) is LMT at +15° longitude.

The hour angle of the apparent sun with respect to the lower branch of the local meridian is called **Local Apparent Time**, **LAT**:

$$LAT[h] = GAT[h] + \frac{Lon[°]}{15}$$

The instant of the apparent sun crossing the local meridian (upper branch) is called **Local Apparent Noon**, **LAN**.

Time measurement by the earth's rotation does not necessarily require the sun as the reference point in the sky. **Greenwich Apparent Sidereal Time**, **GAST**, is a time scale based upon the Greenwich hour angle (upper branch) of the true vernal equinox of date, GHA_{Aries} (see Fig. 3-3).

$$GAST\left[h\right] = \frac{GHA_{Aries}\left[°\right]}{15}$$

The values for GHA_{Aries} tabulated in the Nautical Almanac refer to the true equinox of date.

GAST is easily measured by the Greenwich meridian transit of stars since GAST and the right ascension of the observed star are numerically equal at the moment of meridian transit.

The Greenwich hour angle of the imaginary mean vernal equinox of date (traveling along the celestial equator at a constant rate) defines **Greenwich Mean Sidereal Time**, **GMST**. The difference to GMST is called **equation of the equinoxes**, **EQ**, or **nutation in right ascension**. EQ can be predicted precisely. It varies within approx. ±1s.

$$EQ = GAST - GMST$$

GMST is of some importance for it is the actual basis for UT. Since time measurement by meridian transits of the sun is not accurate enough for many scientific applications, Universal Time is by definition calculated from GMST. This is possible because there is a close correlation between GMST and mean solar time at Greenwich. The time thus obtained is called **UT0**. Applying a small correction (milliseconds) for the effects of **polar motion** (a quasi-circular movement of the earth's crust with respect to the axis of rotation) yields **UT1**, commonly called **UT**.

Due to the earth's revolution around the sun, a mean sidereal day (the time interval between two concecutive meridian transits of the mean equinox) is slightly shorter than a mean solar day:

$$24\,\text{h} \;\; \text{mean sidereal time} = 23\,\text{h}\;56\,\text{m}\;4.090524\,\text{s} \;\; \text{mean solar time}$$

In analogy with LMT and LAT, there is a **Local Mean Sidereal Time, LMST**, and a **Local Apparent Sidereal Time, LAST**:

$$LMST\left[h\right] = GMST\left[h\right] + \frac{Lon\left[°\right]}{15} \qquad LAST\left[h\right] = GAST\left[h\right] + \frac{Lon\left[°\right]}{15}$$

Solar time and sidereal time are both linked to the earth's rotation. The earth's rotational speed, however, decreases slowly (tidal friction) and, moreover, fluctuates in an unpredictable manner due to random movements of matter within the earth's body (magma) and on the surface (water, air). Therefore, neither of both time scales is strictly uniform. Many astronomical applications, however, require a linear time scale. One example is the calculation of ephemerides since the motions of celestial bodies in space are independent of the earth's rotation.

International Atomic Time, **TAI**, is the most precise time standard presently available. It is obtained by statistical analysis of data supplied by a great number of atomic clocks all over the world. Among others, two important time scales are derived from TAI:

Today, civil life is mostly determined by **Coordinated Universal Time, UTC**, which is the basis for time signals broadcast by radio stations, e. g., WWV or WWVH. UTC is controlled by TAI. Due to the varying rotational speed of the earth, UT tends to drift away from UTC. This is undesirable since the cycle of day and night is linked to UT. Therefore, UTC is synchronized to UT, if necessary, by inserting (or omitting) leap seconds at certain times (June 30 and December 31) in order to avoid that the difference, ΔUT, exceeds the specified maximum value of ±0.9 s.

$$UT = UTC + \Delta UT$$

$$UTC = TAI - N$$

N is the cumulative number of leap seconds inserted until now (N = 33 in 2006.0). Due to the occasional leap seconds, UTC is **not** a continuous time scale! Predicted values for ΔUT (= UT1-UTC) are published by the **IERS Rapid Service** [15] on a weekly basis (IERS Bulletin A). The IERS also announces the insertion (or omission) of leap seconds in advance (IERS Bulletins A + C).

Terrestrial Time, TT (formerly called **Terrestrial Dynamical Time**, **TDT**), is another derivative of TAI:

$$TT = TAI + 32.184\,\text{s}$$

TT has replaced **Ephemeris Time**, **ET**. The offset of 32.184 s with respect to TAI is necessary to ensure a seamless continuation of ET. TT is used in astronomy (calculation of ephemerides) and space flight. The difference between TT and UT is called ΔT:

$$\Delta T = TT - UT$$

At the beginning of the year 2006, ΔT was +64.9s. ΔT is of some importance since computer almanacs require TT (TDT) as time argument (programs using UT calculate on the basis of extrapolated ΔT values). A precise long-term prediction of ΔT is impossible. Therefore, computer almanacs using only UT as time argument may become less accurate in the long term. ΔT values for the near future can be calculated with the following formula:

$$\Delta T = 32.184\,\text{s} + (TAI - UTC) - (UT1 - UTC)$$

Like UT1-UTC, TAI-UTC (cumulative number of leap seconds) is published in the IERS Bulletin A.

A final remark on GMT:

The term GMT has become ambigous since it is often used as a synonym for UTC now. Moreover, astronomers used to reckon GMT from the upper branch of the Greenwich meridian until 1925 (the time thus obtained is sometimes called Greenwich Mean Astronomical Time, GMAT). Therefore, the term GMT should be avoided in scientific publications, except when used in a historical context.

The Nautical Almanac

Predicted values for GHA and Dec of sun, moon and the navigational planets with reference to UT are tabulated for each whole hour of the year on the daily pages of the **Nautical Almanac**, **N.A.**, and similar publications [12, 13]. GHA_{Aries} is tabulated in the same manner.

Listing GHA and Dec of all 57 fixed stars used in navigation for each whole hour of the year would require too much space. Since declinations and sidereal hour angles of stars change only slowly, tabulated values for periods of 3 days are accurate enough for navigational purposes.

GHA and Dec for each second of the year are obtained using the **interpolation tables** at the end of the N.A. (printed on tinted paper), as explained in the following directions:

1.

We note the exact time of observation (UT), determined with a chronometer, for each celestial body. If UT is not available, we can use UTC. The resulting error is tolerable in most cases.

2.

We look up the day of observation in the N.A. (two pages cover a period of three days).

3.

We go to the nearest whole hour preceding the time of observation and note GHA and Dec of the observed body. In case of a fixed star, we form the sum of GHA Aries and the SHA of the star, and note the tabulated declination. When observing planets, we note the ***v*** and ***d*** factors given at the bottom of the appropriate column. For the moon, we take *v* and *d* for the nearest whole hour preceding the time of observation.

The quantity *v* is necessary to apply an additional correction to the following interpolation of the GHA of moon and planets. It is not required for stars. The sun does not require a *v* factor since the correction has been incorporated in the tabulated values for the sun's GHA.

The quantity *d*, which is negligible for stars, is the rate of change of Dec, measured in arcminutes per hour. It is needed for the interpolation of Dec. The sign of *d* is critical!

4.

We look up the minute of observation in the interpolation tables (1 page for each 2 minutes of the hour), go to the second of observation, and note the **increment** from the respective column.

We enter one of the three columns to the right of the increment columns with the *v* and *d* factors and note the corresponding **corr**(ection) values (*v*-corr and *d*-corr).

The sign of *d*-corr depends on the trend of declination at the time of observation. It is positive if Dec at the whole hour following the observation is greater than Dec at the whole hour preceding the observation. Otherwise it is negative.

v -corr is negative for Venus. Otherwise, it is always positive.

5.

We form the sum of Dec and *d*-corr (if applicable).

We form the sum of GHA (or GHA Aries and SHA of star), increment, and *v*-corr (if applicable).
SHA values tabulated in the Nautical Almanac refer to the true vernal equinox of date.

Interactive Computer Ephemeris

Interactive Computer Ephemeris, **ICE**, is a computer almanac developed by the U.S. Naval Observatory (successor of the **Floppy Almanac**).

ICE is FREEWARE (no longer supported by USNO), compact, easy to use, and provides a vast quantity of accurate astronomical data for a time span of almost 250 (!) years. In spite of the old design (DOS program), ICE is an extremely useful tool for navigators and astronomers.

Among many other features, ICE calculates GHA and Dec for a given body and time as well as altitude and azimuth of the body for an **assumed position** (see chapter 4) and sextant altitude corrections. Since the navigation data are as accurate as those tabulated in the **Nautical Almanac** (approx. 0.1'), the program makes an adequate alternative, although a printed almanac (and sight reduction tables) should be kept as a backup in case of a computer failure. The following instructions refer to the final version (0.51). Only program features relevant to navigation are explained.

1. Installation

Copy the program files to a chosen directory on the hard drive, floppy disk, USB stick, or similar storage device.

2. Getting Started

Go to the program directory and enter "ice". The main menu appears.
Use the function keys F1 to F10 to navigate through the submenus. The program is more or less self-explanatory.
Go to the submenu INITIAL VALUES (F1). Follow the directions on the screen to enter date and time of observation (F1), assumed latitude (F2), assumed longitude (F3), and your local time zone (F6). Assumed latitude and longitude define your assumed position.

Use the correct data format, as shown on the screen (decimal format for latitude and longitude). After entering the above data, press F7 to accept the values displayed. To change the default values permanently, edit the file ice.dft with a text editor (after making a backup copy) and make the desired changes. Do not change the data format. The numbers have to be in columns 21-40. An output file can be created to store calculated data. Go to the submenu FILE OUTPUT (F2) and enter a chosen file name, e.g., OUTPUT.TXT.

3. Calculation of Navigational Data

From the main menu, go to the submenu NAVIGATION (F7). Enter the name of the body. The program displays GHA and Dec of the body, GHA and Dec of the sun (if visible), and GHA of the vernal equinox for the date and time (UT) stored in INITIAL VALUES.
Hc (computed altitude) and Zn (azimuth) mark the apparent position of the body as observed from the assumed position.

Approximate altitude corrections (refraction, SD, PA), based upon Hc, are also displayed (for lower limb of body). The semidiameter of the moon includes augmentation. The coordinates calculated for Venus and Mars do **not** include phase correction. Therefore, the upper or lower limb (if visible) should be observed. ΔT is TDT-UT, the predicted difference between **terrestrial dynamical time** and UT for the given date. The ΔT value for 2006.0 predicted by ICE is 67.8s, the actual value is 64.9s (see below).

Horizontal parallax and semidiameter of a body can be extracted, if required, from the submenu POSITIONS (F3). Choose APPARENT GEOCENTRIC POSITIONS (F1) and enter the name of the body (sun, moon, planets). The last column shows the distance of the center of the body from the center of the earth, measured in astronomical units (1 AU = $149.6 \cdot 10^6$ km). HP and SD are calculated as follows:

$$HP = \arcsin \frac{r_E[km]}{distance[km]} \qquad SD = \arcsin \frac{r_B[km]}{distance[km]}$$

r_E is the equatorial radius of the earth (6378 km). r_B is the radius of the body (Sun: 696260 km, Moon: 1378 km, Venus: 6052 km, Mars: 3397 km, Jupiter: 71398 km, Saturn: 60268 km).

The apparent geocentric positions refer to TDT, but the difference between TDT and UT has no significant effect on HP and SD.

To calculate times of rising and setting of a body, go to the submenu RISE & SET TIMES (F6) and enter the name of the body. The columns on the right display the time of rising, meridian transit, and setting for the assumed location (UT+xh, according to the time zone specified).

The increasing error of ΔT values predicted by ICE may lead to reduced precision when calculating navigation data for the future. The coordinates of the moon are particularly sensitive to errors of ΔT. Unfortunately, ICE has no option for editing and modifying ΔT. The high-precision part of ICE, however, is not affected since TDT is the time argument. To extract GHA and Dec, the following procedure is recommended:

1. Compute GAST using SIDEREAL TIME (F5). The time argument is UT.

2. Edit date and time at INITIAL VALUES (F1). Now, the time argument is UT+ΔT. Compute RA and Dec using POSITIONS (F3) and APPARENT GEOCENTRIC POSITIONS (F1).

3. Use the following formula to calculate GHA from GAST and RA (RA refers to the true vernal equinox of date):

$$GHA[°] = 15 \cdot (GAST[h] + 24h - RA[h])$$

Add or subtract 360° if necessary.

High-precision GHA and Dec values thus obtained can be used as an internal standard to check medium-precision data obtained through NAVIGATION (F7).

Chapter 4

Finding One's Position (Sight Reduction)

Lines of Position

Any geometrical or physical line passing through the observer's (still unknown) position and accessible through measurement or observation is called a **line of position** or **position line**, **LOP**. Examples are circles of equal altitude, meridians, parallels of latitude, bearing lines (compass bearings) of terrestrial objects, coastlines, rivers, roads, railroad tracks, power lines, etc. **A single position line indicates an infinite series of possible positions.** The observer's actual position is marked by the point of intersection of at least two position lines, regardless of their nature. A position thus found is called **fix** in navigator's language. The concept of the position line is essential to modern navigation.

Sight Reduction

Finding a line of position by observation of a celestial object is called **sight reduction**. Although some background in mathematics is required to comprehend the process completely, knowing the basic concepts and a few equations is sufficient for most practical applications. The geometrical background (law of cosines, navigational triangle) is given in chapter 10 and 11. In the following, we will discuss the semi-graphic methods developed by *Sumner* and *St. Hilaire*. Both methods require relatively simple calculations only and enable the navigator to plot lines of position on a navigation chart or **plotting sheet** (chapter 13).

Knowing altitude and GP of a body, we also know the radius of the corresponding circle of equal altitude (our line of position) and the location of its center. As mentioned in chapter 1 already, plotting circles of equal altitude on a chart is usually impossible due to their large dimensions and the distortions caused by map projection. However, *Sumner* and *St. Hilaire* showed that only a short arc of each circle of equal altitude is needed to find one's position. Since this arc is comparatively short, it can be replaced with a secant or tangent a of the circle.

Local Meridian, Local Hour Angle and Meridian Angle

A meridian passing through a given position is called **local meridian**. In celestial navigation, the angle formed by the hour circle of the observed body (upper branch) and the local meridian (upper branch) plays a fundamental role. On the analogy of the Greenwich hour angle, we can measure this angle westward from the local meridian (0°...+360°). In this case, the angle is called **local hour angle**, **LHA**. It is also possible to measure the angle westward (0°...+180°) or eastward (0°...–180°) from the local meridian in wich case it is called **meridian angle**, **t**. In most navigational formulas, LHA and t can be substituted for each other since the trigonometric functions return the same results for either of both angles. For example, the cosine of +315° is the same as the cosine of –45°.

LHA as well as t is the algebraic sum of the Greenwich hour angle of the body, GHA, and the observer's longitude, Lon. To make sure that the obtained angle is in the desired range, the following rules have to be applied when forming the sum of GHA and Lon:

$$LHA = \begin{cases} GHA + Lon & \text{if} \quad 0° < GHA + Lon < 360° \\ GHA + Lon + 360° & \text{if} \quad GHA + Lon < 0° \\ GHA + Lon - 360° & \text{if} \quad GHA + Lon > 360° \end{cases}$$

$$t = \begin{cases} GHA + Lon & \text{if} \quad GHA + Lon < 180° \\ GHA + Lon - 360° & \text{if} \quad GHA + Lon > 180° \end{cases}$$

In all calculations, the sign of Lon and t, respectively, has to be observed carefully. The sign convention is:

Eastern longitude: **positive**
Western longitude: **negative**

Eastern meridian angle: **negative**
Western meridian angle: **positive**

For reasons of symmetry, we will refer to the meridian angle in the following considerations (a body has the same altitude with the meridian angle +t and −t, respectively), although the local hour angle would lead to the same results.

Fig. 4-1 illustrates the various angles involved in the sight reduction process.

Sumner's Method

In December 1837, *Thomas H Sumner*, an American sea captain, was on a voyage from South Carolina to Greenock, Scotland. When approaching *St. George's Channel* between Ireland and Wales, he managed to measure a single altitude of the sun after a longer period of bad weather. Using the time sight formula (see chapter 6), he calculated a longitude from his estimated latitude. Since he was doubtful about his estimate, he repeated his calculations with two slightly different latitudes. To his surprise, the three points thus obtained were on a straight line. Accidentally, the line passed through the position of a light house off the coast of Wales (*Small's Light*). By intuition, Sumner steered his ship along this line and soon after, *Small's Light* came in sight. Sumner concluded that he had found a "line of equal altitude". The publication of his method in 1843 marked the beginning of "modern" celestial navigation [18]. Although rarely used today, it is still an interesting alternative. It is easy to comprehend and the calculations to be done are extremely simple.

Fig. 4-2 illustrates the points where a circle of equal altitude intersects two chosen parallals of latitude.

An observer being between Lat_1 and Lat_2 is either on the arc A-B or on the arc C-D. With a rough estimate of his longitude, the observer can easily find on which of both arcs he is, for example, A-B. The arc thus found is the relevant part of his line of position, the other arc is discarded. We can approximate the line of position by drawing a straight line through A and B which is a secant of the circle of equal altitude. This secant is called **Sumner line**. Before plotting the Sumner line on our chart, we have to find the longitude of each point of intersection, A, B, C, and D.

Procedure:

1.

We choose a parallel of latitude (Lat_1) north of our estimated latitude. Preferably, Lat_1 should be marked by the nearest horizontal grid line on our chart or plotting sheet.

2.

From Lat_1, Dec, and the observed altitude, Ho, we calculate the meridian angle, t, using the following formula:

$$t = \pm \arccos \frac{\sin Ho - \sin Lat \cdot \sin Dec}{\cos Lat \cdot \cos Dec}$$

The equation is derived from the navigational triangle (chapter 10 & 11). It has two solutions, +t and –t, since the cosine of +t equals the cosine of –t. Geometrically, this corresponds with the fact that the circle of equal altitude intersects the parallel of latitude at two points. Using the following formulas and rules, we obtain the longitudes of these points of intersection, Lon and Lon':

$$Lon = t - GHA$$

$$Lon' = 360° - t - GHA$$

$$\text{If } Lon < -180° \quad \rightarrow \quad Lon + 360°$$
$$\text{If } Lon' < -180° \quad \rightarrow \quad Lon' + 360°$$
$$\text{If } Lon' > +180° \quad \rightarrow \quad Lon' - 360°$$

Comparing the longitudes thus obtained with our estimate, we select the most probable longitude and discard the other one. This method of finding one's longitude is called **time sight** (see chapter 6).

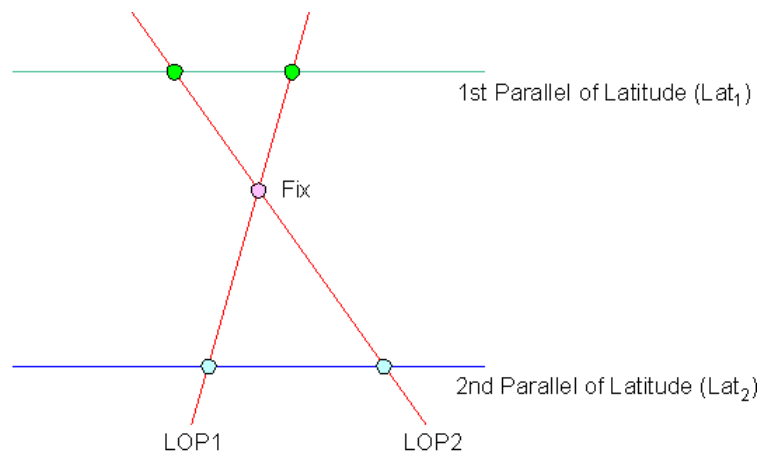
3.

We chose a parallel of latitude (Lat_2) south of our estimated latitude. The difference between Lat_1 and Lat_2 should not exceed 1 or 2 degrees. We repeat steps 1 and 2 with the second latitude, Lat_2.

4.

On our plotting sheet, we mark each remaining longitude on the corresponding parallel and plot the Sumner line through the points thus located (LOP1).

To obtain a fix, we repeat steps 1 through 4 with the same parallels and the declination and observed altitude of a second body. The point where the Sumner line thus obtained, LOP2, intersects LOP1 is our fix (*Fig. 4-3*).

Fig. 4-3

If we have only a very rough estimate of our latitude, the point of intersection may be slightly outside the interval defined by both parallels. Nevertheless, the fix is correct. A fix obtained with *Sumner's* method has a small error caused by neglecting the curvature of the circles of equal altitude. We can improve the fix by iteration. In this case, we choose a new pair of assumed latitudes, nearer to the fix, and repeat the procedure. Ideally, the horizontal distance between both bodies should be 90° (30°...150° is tolerable). Otherwise, the fix would become indistinct. Further, neither of the bodies should be near the local meridian (see time sight, chapter 6). *Sumner's* method has the (small) advantage that no protractor is needed to plot lines of position.

The Intercept Method

This procedure was developed by the French navy officer *St. Hilaire* and others and was first published in 1875. After that, it gradually became the standard for sight reduction since it avoids some of the restrictions of *Sumner's* method. Although the background is more complicated than with *Sumner's* method, the practical application is very convenient.

Theory:

For any given position of the observer, the altitude of a celestial body, reduced to the celestial horizon, is **solely** a function of the observer's latitude, the declination of the body, and the meridian angle (or local hour angle). The **altitude formula** is obtained by applying the law of cosine for sides to the navigational triangle (see chapter 10 & 11):

$$H = \arcsin\left(\sin Lat \cdot \sin Dec + \cos Lat \cdot \cos Dec \cdot \cos t\right)$$

We choose an arbitrary point in the vicinity of our estimated position, preferably the nearest point where two grid lines on the chart intersect. This point is called **assumed position**, **AP** (*Fig. 4-2*). Using the above formula, we calculate the altitude of the body resulting from Lat_{AP} and Lon_{AP}, the geographic coordinates of AP. The altitude thus obtained is called **computed** or **calculated altitude**, **Hc**.

Usually, Hc will slightly differ from the actually **observed altitude**, **Ho** (see chapter 2). The difference, ΔH, is called **intercept**.

$$\Delta H = Ho - Hc$$

Ideally, Ho and Hc are identical if the observer is at AP.

In the following, we will discuss which possible positions of the observer would result in the same intercept, ΔH. For this purpose, we assume that the intercept is an infinitesimal quantity and denote it by dH. The general formula is:

$$dH = \frac{\partial H}{\partial Lat} \cdot dLat + \frac{\partial H}{\partial t} \cdot dt$$

This differential equation has an infinite number of solutions. Since dH and both differential coefficients are constant, it can be reduced to an equation of the general form:

$$dLat = a + b \cdot dt$$

Thus, the graph is a straight line, and it is sufficient to dicuss two special cases, dt=0 and dLat=0, respectively.

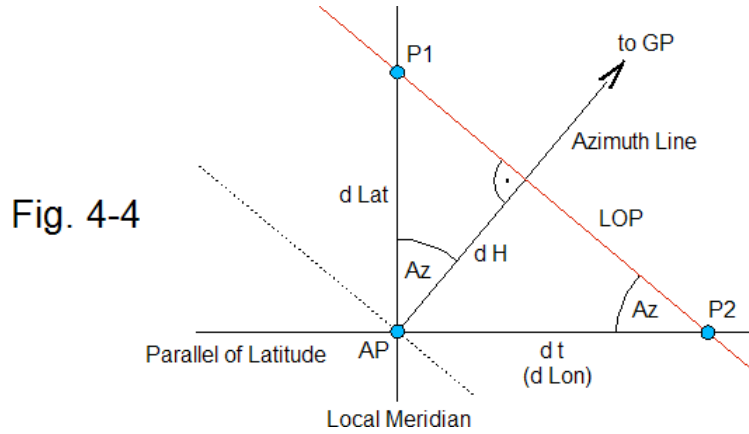
In the first case, the observer is on the same meridian as AP, and dH is solely caused by a small variation of latitude, dLat, whereas t is constant (dt = 0). We differentiate the altitude formula with respect to Lat:

$$\sin H = \sin Lat \cdot \sin Dec + \cos Lat \cdot \cos Dec \cdot \cos t$$

$$d\left(\sin H\right) = \left(\cos Lat \cdot \sin Dec - \sin Lat \cdot \cos Dec \cdot \cos t\right) \cdot dLat$$

$$\cos H \cdot dH = \left(\cos Lat \cdot \sin Dec - \sin Lat \cdot \cos Dec \cdot \cos t\right) \cdot dLat$$

$$dLat = \frac{\cos H}{\cos Lat \cdot \sin Dec - \sin Lat \cdot \cos Dec \cdot \cos t} \cdot dH$$

Adding dLat to Lat_{AP}, we obtain the point P1, as illustrated in *Fig.4-4*. P1 is on the observer's circle of equal altitude.

In the second case, the observer is on the same parallel of latitude as AP, and dH is solely caused by a small change of the meridian angle, dt, whereas Lat is constant (dLat=0). We differentiate the altitude formula with respect to t:

$$\sin H = \sin Lat \cdot \sin Dec + \cos Lat \cdot \cos Dec \cdot \cos t$$

$$d(\sin H) = -\cos Lat \cdot \cos Dec \cdot \sin t \cdot dt$$

$$\cos H \cdot dH = -\cos Lat \cdot \cos Dec \cdot \sin t \cdot dt$$

$$dt = -\frac{\cos H}{\cos Lat \cdot \cos Dec \cdot \sin t} \cdot dH$$

Adding dt (corresponding with an equal change of longitude, dLon) to Lon_{AP}, we obtain the point P2 which, too, is on the observer's circle of equal altitude. Thus, we would measure Ho at P1 and P2, respectively. Knowing P1 and P2, we can plot a straight line passing through these positions. This line is a tangent of the circle of equal altitude and is our **line of position**, **LOP**. The great circle passing through AP and GP is represented by a straight line perpendicular to the line of position. The arc between AP and GP is the radius of the circle of equal altitude. The distance between AP and the point where this line, called **azimuth line**, intersects the line of position is the intercept, dH. The angle formed by the azimuth line and the local meridian of AP is called **azimuth angle**, **Az**. The same angle is formed by the line of position and the parallel of latitude passing through AP (*Fig. 4-4*).

There are several ways to obtain Az and the true azimuth, Az_N, from the right (plane) triangle formed by AP, P1, and P2:

1. Time-altitude azimuth:

$$\cos Az = \frac{dH}{dLat} = \frac{\cos Lat \cdot \sin Dec - \sin Lat \cdot \cos Dec \cdot \cos t}{\cos H}$$

$$Az = \arccos\left(\frac{\cos Lat \cdot \sin Dec - \sin Lat \cdot \cos Dec \cdot \cos t}{\cos H}\right)$$

Az is not necessarily identical with the true azimuth, Az_N, since the arccos function returns angles between 0° and +180°, whereas Az_N is measured from 0° to +360°.

To obtain Az_N, we have to apply the following rules when using the formula for time-altitude azimuth:

$$Az_N = \begin{cases} Az & \text{if} \quad t < 0° \quad (180° < LHA < 360°) \\ 360° - Az & \text{if} \quad t > 0° \quad (0° < LHA < 180°) \end{cases}$$

2. Time azimuth:

$$\tan Az = \frac{d\,Lat}{\cos Lat \cdot dt} = \frac{\sin t}{\sin Lat \cdot \cos t - \cos Lat \cdot \tan Dec}$$

The factor cos Lat is the relative circumference of the parallel of latitude going through AP (equator = 1).

$$Az = \arctan \frac{\sin t}{\sin Lat \cdot \cos t - \cos Lat \cdot \tan Dec}$$

The time azimuth formula does not require the altitude. Since the arctan function returns angles between -90° and +90°, a different set of rules is required to obtain Az_N:

$$Az_N = \begin{cases} Az & \text{if} \quad \text{numerator} < 0 \quad \text{AND} \quad \text{denominator} < 0 \\ Az + 360° & \text{if} \quad \text{numerator} > 0 \quad \text{AND} \quad \text{denominator} < 0 \\ Az + 180° & \text{if} \quad \text{denominator} > 0 \end{cases}$$

3. Altitude azimuth:

This formula is directly derived from the navigational triangle (cosine law, see chapter 10 & 11) without using differential calculus.

$$\cos Az = \frac{\sin Dec - \sin H \cdot \sin Lat}{\cos H \cdot \cos Lat}$$

$$Az = \arccos \frac{\sin Dec - \sin Hc \cdot \sin Lat}{\cos Hc \cdot \cos Lat}$$

As with the formula for time-altitude azimuth, Az_N is obtained through these rules:

$$Az_N = \begin{cases} Az & \text{if} \quad t < 0 \quad (180° < LHA < 360°) \\ 360° - Az & \text{if} \quad t > 0 \quad (0° < LHA < 180°) \end{cases}$$

In contrast to dH, ΔH is a measurable quantity, and the position line is curved. *Fig. 4-5* shows a macroscopic view of the line of position, the azimuth line, and the circles of equal altitude.

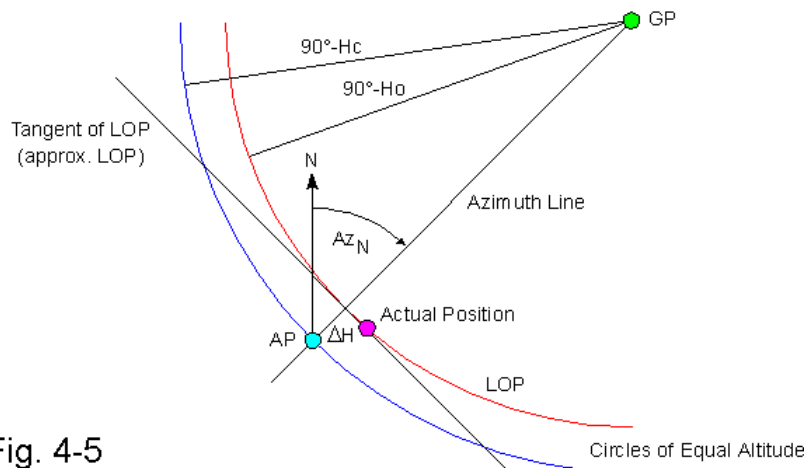



Fig. 4-5

Procedure:

Although the theory of the intercept method looks complicated, its practical application is very simple and does not require any background in differential calculus. The procedure comprises the following steps:

1.

We choose an **assumed position**, **AP**, near to our **estimated position**. Preferably, AP should be defined by an integer number of degrees for Lat_{AP} and Lon_{AP}, respectively, depending on the scale of the chart. Instead of AP, our estimated position itself may be used. Plotting lines of position, however, is more convenient when putting AP on the point of intersection of two grid lines.

2.

We calculate the meridian angle, t_{AP}, (or local hour angle, LHA_{AP}) from GHA and Lon_{AP}, as stated above.

3.

We calculate the altitude of the observed body as a function of Lat_{AP}, t_{AP}, and Dec (computed altitude):

$$Hc = \arcsin\left(\sin Lat_{AP} \cdot \sin Dec + \cos Lat_{AP} \cdot \cos Dec \cdot \cos t_{AP}\right)$$

4.

Using one of the azimuth formulas stated above, we calculate the **true azimuth** of the body, $\mathbf{Az_N}$, from Hc, Lat_{AP}, t_{AP}, and Dec, for example:

$$Az = \arccos \frac{\sin Dec - \sin Hc \cdot \sin Lat_{AP}}{\cos Hc \cdot \cos Lat_{AP}}$$

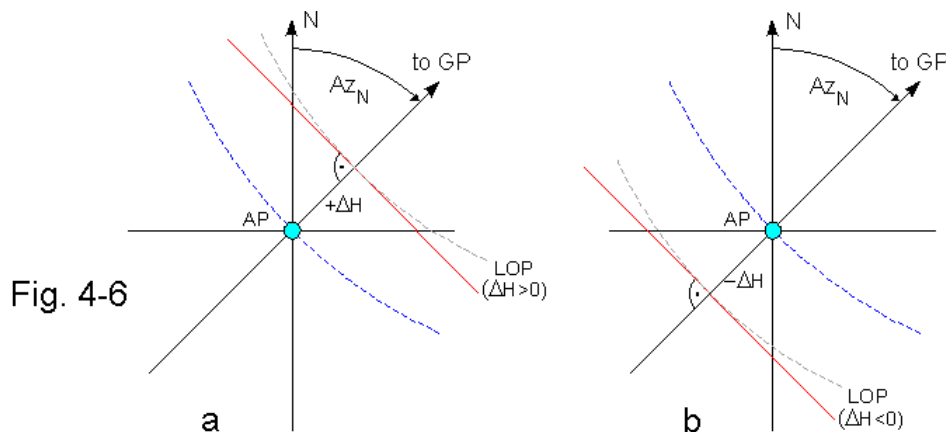
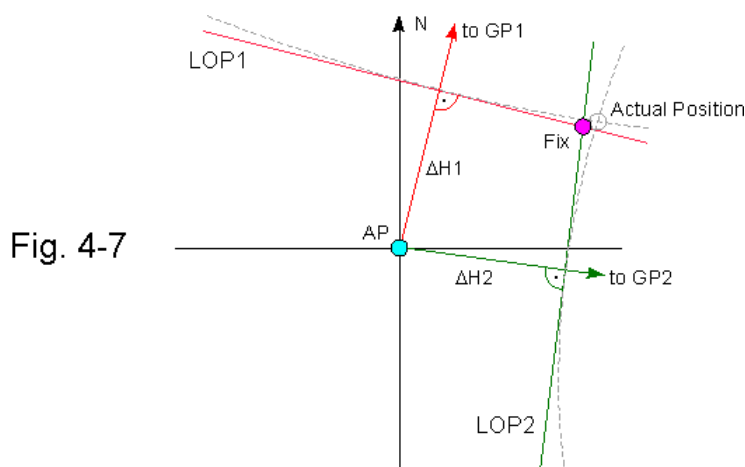
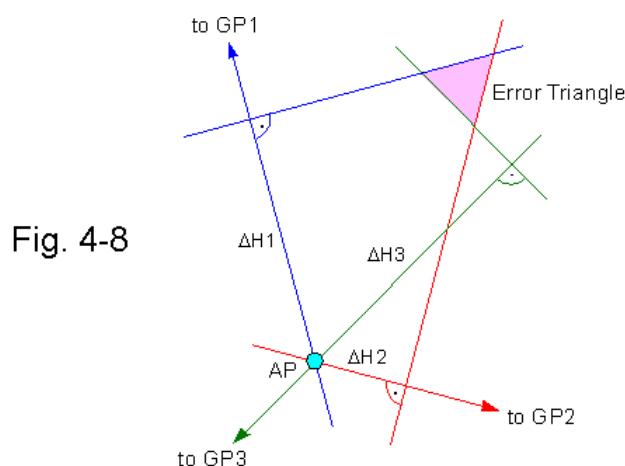
$$Az_N = \begin{cases} Az & \text{if} \quad t < 0 \quad (180° < LHA < 360°) \\ 360° - Az & \text{if} \quad t > 0 \quad (0° < LHA < 180°) \end{cases}$$

5.

We calculate the **intercept**, **ΔH**, the difference between observed altitude, Ho (chapter 2), and computed altitude, Hc. The intercept, which is directly proportional to the difference between the radii of the corresponding circles of equal altitude, is usually expressed in nautical miles:

$$\Delta H\,[nm] = 60 \cdot \left(Ho\,[°] - Hc\,[°] \right)$$

6.

On the chart, we draw a suitable length of the azimuth line through AP (*Fig. 4-6*). On this line, we measure the intercept, ΔH, from AP (towards GP if ΔH>0, away from GP if ΔH<0) and draw a perpendicular through the point thus located. This perpendicular is our approximate line of position.

To obtain our position, we need at least one more line of position. We repeat the procedure with altitude and GP of a second celestial body or of the same body at a different time of observation (*Fig. 4-4*). The point where both position lines (tangents) intersect is our fix. The second observation does not necessarily require the same AP to be used.

As mentioned above, the intercept method ignores the curvature of the actual LoP's. Therefore, the obtained fix is not our exact position but an **improved position** (compared with AP). The residual error remains tolerable as long as the radii of the circles of equal altitude are great enough and AP is not too far from the actual position (see chapter 16). The geometric error inherent to the intercept method can be decreased by **iteration**, i.e., substituting the obtained fix for AP and repeating the calculations (same altitudes and GP's). This will result in a more accurate position. If necessary, we can reiterate the procedure until the obtained position remains virtually constant. Since an estimated position is usually nearer to our true position than an assumed position, the latter may require a greater number of iterations. Accuracy is also improved by observing three bodies instead of two. Theoretically, the position lines should intersect each other at a single point. Since no observation is entirely free of errors, we will usually obtain three points of intersection forming an **error triangle** (*Fig. 4-8*).

Area and shape of the triangle give us a rough estimate of the quality of our observations (see chapter 16). Our **most probable position**, **MPP**, is near the center of the inscribed circle of the error triangle (the point where the bisectors of the three angles of the error triangle meet).

When observing more than three bodies, the resulting position lines will form the corresponding polygons.

Direct Computation

If we do not want to plot lines of position to determine our fix, we can calculate the most probable position directly from an unlimited number of observations, n (n > 1). The Nautical Almanac provides an averaging procedure.

First, the auxiliary quantities A, B, C, D, E, and G have to be calculated:

$$A = \sum_{i=1}^{n} \cos^2 Az_i \qquad B = \sum_{i=1}^{n} \sin Az_i \cdot \cos Az_i \qquad C = \sum_{i=1}^{n} \sin^2 Az_i$$

$$D = \sum_{i=1}^{n} (\Delta H)_i \cdot \cos Az_i \qquad E = \sum_{i=1}^{n} (\Delta H)_i \cdot \sin Az_i \qquad G = A \cdot C - B^2$$

In these formulas, Az_i denotes the *true* azimuth of the respective body. The ΔH values are measured in degrees (same unit as Lon and Lat). The geographic coordinates of the observer's MPP are then obtained as follows:

$$Lon = Lon_{AP} + \frac{A \cdot E - B \cdot D}{G \cdot \cos Lat_{AP}} \qquad Lat = Lat_{AP} + \frac{C \cdot D - B \cdot E}{G}$$

The method does not correct for the geometric errors inherent to the intercept method. These are eliminated, if necessary, by iteration. For this purpose, we substitute the calculated MPP for AP. For each body, we calculate new values for t (or LHA), Hc, ΔH, and Az_N. With these values, we recalculate A, B, C, D, E, G, Lon, and Lat.

Repeating this procedure, the resulting positions will converge rapidly. In the majority of cases, one or two iterations will be sufficient, depending on the distance between AP and the true position.

Combining Different Lines of Position

Since the point of intersection of any two position lines, regardless of their nature, marks the observer's geographic position, one celestial LOP may suffice to find one's position if another LOP of a different kind is available.

In the desert, for instance, we can determine our current position by finding the point on the map where a position line obtained by observation of a celestial object intersects the dirt road we are using (*Fig. 4-9*).

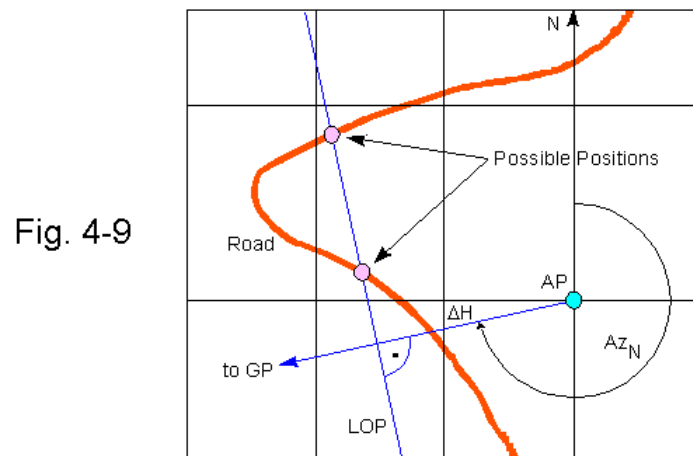
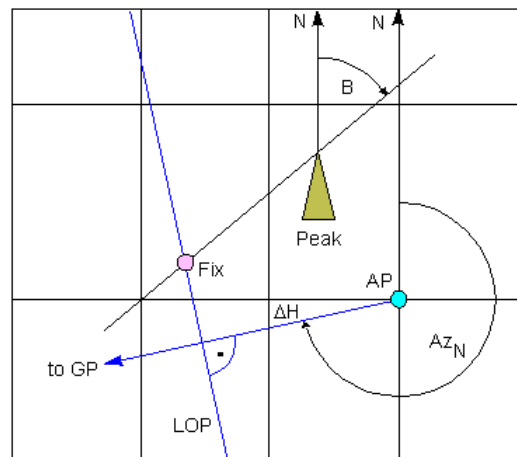



Fig. 4-9

We can as well find our position by combining our celestial LOP with the bearing line of a distant mountain peak or any other prominent landmark (*Fig. 4-10*). B is the compass bearing of the terrestrial object (corrected for magnetic declination).

Fig. 4-10

Both examples demonstrate the versatility of position line navigation.

Chapter 5

Finding the Position of a Moving Vessel

Celestial navigation on a moving vessel requires special measures to correct for the change in position between different observations unless the latter are performed in rapid succession or simultaneously, e. g., by a second observer.

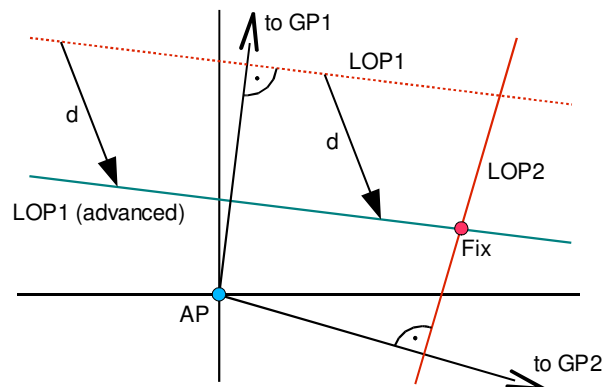
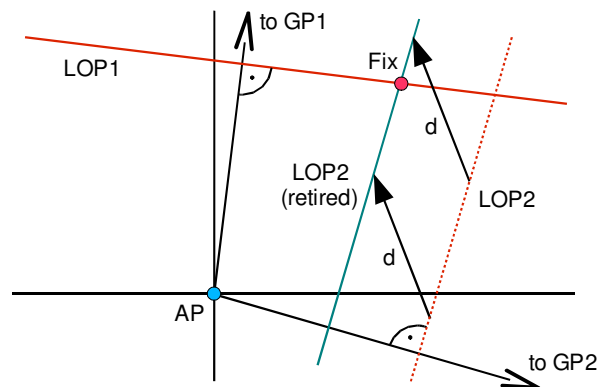
If the navigator knows the speed of the vessel, v, and the true course, C (the angle formed by the motion vector and the local meridian), position line navigation provides a simple graphic solution.

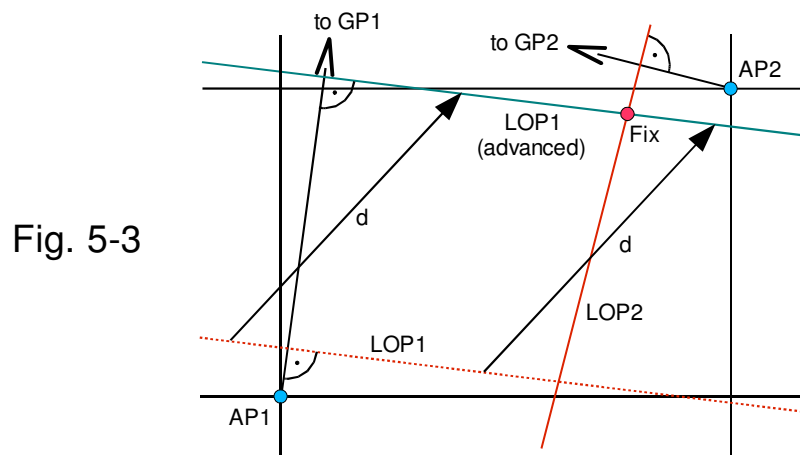
Assuming that we make our first observation at the time T_1 and our second observation at T_2, the distance, d, traveled during the time interval T_2-T_1 is

$$d[nm] = (T_2[h] - T_1[h]) \cdot v[kn]$$

1 kn (knot) = 1 nm/h

Although we have no knowledge of our absolute position yet, we know our second position relative to the first one.

To find the absolute position, we plot both position lines in the usual manner, as illustrated in chapter 4. Then, we choose an arbitrary point of the first position line (resulting from the observation at T_1) and advance this point according to the motion vector defined by d and C. Next, we draw a parallel of the first position line through the point thus located. The point where this **advanced position line** intersects the second line of position (resulting from the observation at T_2) marks our position at $\mathbf{T_2}$. A position obtained in this fashion is called **running fix** (*Fig. 5-1*).

In a similar manner, we can obtain our position at $\mathbf{T_1}$ by **retiring** the second position line (*Fig. 5-2*).

Terrestrial lines of position may be advanced or retired in the same way as astronomical position lines.

It is also possible to choose two different assumed positions. AP1 should be close to the estimated position at T_1, AP2 close to the estimated position at T_2 (*Fig. 5-3*).

A running fix is not as accurate as a stationary fix. For one thing, course and speed over ground can only be estimated since the effects of current and wind (drift) are not exactly known in most cases.

Further, there is a geometrical error inherent to the method. The latter is based on the assumption that each point of the circle of equal altitude, representing a possible position of the vessel, travels the same distance, d, along the **rhumb line** (see chapter 12) defined by the course, C. The result of such an operation, however, is a distorted circle. Consequently, the advanced (or retired) LOP is not exactly parallel to the original LOP. The resulting position error usually increases as the distance, d, increases [19]. The procedure gives fairly accurate results when the distance traveled between the observations is smaller than approx. 50 nm.

Chapter 6

Determination of Latitude and Longitude, Direct Calculation of Position

Latitude by Polaris

The observed altitude of a star being vertically above the geographic north pole would be numerically equal to the latitude of the observer (*Fig. 6-1*).

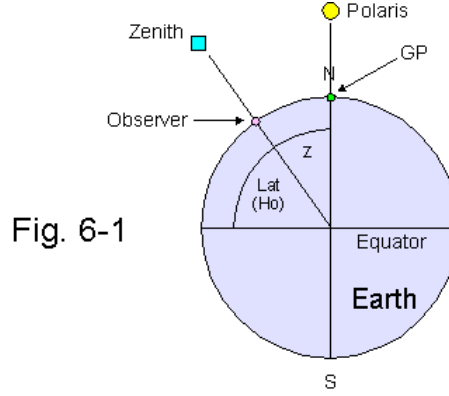



Fig. 6-1

This is nearly the case with the pole star (Polaris). However, since there is a measurable angular distance between Polaris and the polar axis of the earth (presently ca. 1°), the altitude of Polaris is a function of the local hour angle. The altitude of Polaris is also affected by nutation. To obtain the accurate latitude, several corrections have to be applied:

$$Lat = Ho - 1° + a_0 + a_1 + a_2$$

The corrections a_0, a_1, and a_2 depend on LHA_{Aries}, the observer's estimated latitude, and the number of the month. They are given in the Polaris Tables of the Nautical Almanac [12]. To extract the data, the observer has to know his approximate position and the approximate time.

When using a computer almanac instead of the N. A., we can calculate Lat with the following simple procedure. Lat_E is our estimated latitude, Dec is the declination of Polaris, and t is the meridian angle of Polaris (calculated from GHA and our estimated longitude). Hc is the computed altitude, Ho is the observed altitude (see chapter 4).

$$Hc = \arcsin\left(\sin Lat_E \cdot \sin Dec + \cos Lat_E \cdot \cos Dec \cdot \cos t\right)$$

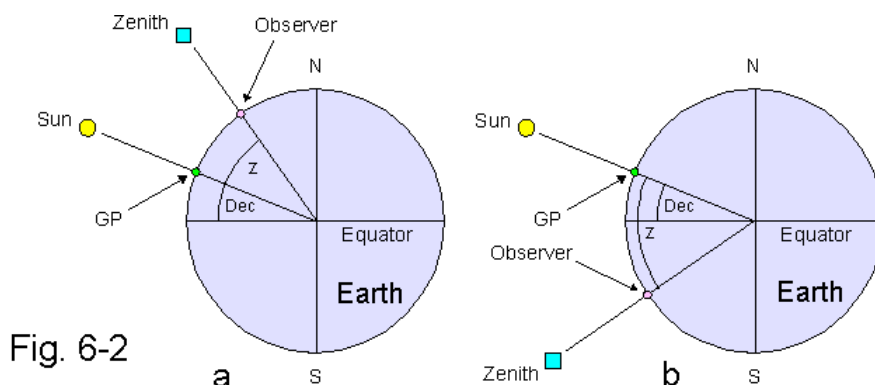
$$\Delta H = Ho - Hc$$

Adding the altitude difference, ΔH, to the estimated latitude, we obtain the improved latitude:

$$Lat \approx Lat_E + \Delta H$$

The error of Lat is smaller than 0.1' when Lat_E is smaller than 70° and when the error of Lat_E is smaller than 2°, provided the exact longitude is known.. In polar regions, the algorithm becomes less accurate. However, the result can be improved by iteration (substituting Lat for Lat_E and repeating the calculation). Latitudes greater than 85° should be avoided because a greater number of iterations might be necassary. The method may lead to erratic results when the observer is close to the north pole ($Lat_E \approx Dec_{Polaris}$). An error in Lat resulting from an error in longitude is **not** decreased by iteration. However, this error is always smaller than 1' when the error in longitude is smaller than 1°.

Noon Latitude (Latitude by Maximum Altitude)

This is a very simple method enabling the observer to determine latitude by measuring the maximum altitude of the sun (or any other object). No accurate time measurement is required. The altitude of the sun passes through a flat maximum **approximately** (see noon longitude) at the moment of upper meridian passage (**local apparent noon, LAN**) when the GP of the sun has the same longitude as the observer and is either north or south of him, depending on the declination of the sun and observer's geographic latitude. The observer's latitude is easily calculated by forming the algebraic sum or difference of the declination and observed zenith distance z (90°-Ho) of the sun, depending on whether the sun is north or south of the observer (*Fig. 6-2*).

1. Sun south of observer (*Fig. 6-2a*): $Lat = Dec + z = Dec - Ho + 90°$

2. Sun north of observer (*Fig. 6-2b*): $Lat = Dec - z = Dec + Ho - 90°$

Northern declination is positive, southern negative.

Before starting the observations, we need a rough estimate of our current longitude to know the time (GMT) of meridian transit. We look up the time of Greenwich meridian transit of the sun on the daily page of the Nautical Almanac and add 4 minutes for each degree of western longitude or subtract 4 minutes for each degree of eastern longitude. To determine the maximum altitude, we start observing the sun approximately 15 minutes before meridian transit. We follow the increasing altitude of the sun with the sextant, note the maximum altitude when the sun starts descending again, and apply the usual corrections.

We look up the declination of the sun at the approximate time (GMT) of local meridian passage on the daily page of the Nautical Almanac and apply the appropriate formula.

Historically, noon latitude and latitude by Polaris are among the oldest methods of celestial navigation.

Ex-Meridian Sight

Sometimes, it may be impossible to measure the maximum altitude of the sun. For example, the sun may be obscured by a cloud at this instant. If we have a chance to measure the altitude of the sun a few minutes before or after meridian transit, we are still able to find our exact latitude by reducing the observed altitude to the meridian altitude, provided we know our exact longitude (see below) and approximate latitude. The method is similar to the one used with the pole star. First, we need the time (UT) of local meridian transit (eastern longitude is positive, western longitude negative):

$$T_{Transit}[h] = 12 - EoT[h] - \frac{Lon[°]}{15}$$

The meridian angle of the sun, t, is calculated from the time of observation (GMT):

$$t[°] = 15 \cdot \left(T_{Observation}[h] - T_{Transit}[h] \right)$$

Starting with our estimated Latitude, Lat_E, we calculate the altitude of the sun at the time of observation. We use the altitude formula from chapter 4:

$$Hc = \arcsin\left(\sin Lat_E \cdot \sin Dec + \cos Lat_E \cdot \cos Dec \cdot \cos t\right)$$

Dec refers to the time of observation. We calculate the difference between observed and calculated altitude:

$$\Delta H = Ho - Hc$$

We calculate an improved latitude, $Lat_{improved}$:

$$Lat_{improved} \approx Lat_E \pm \Delta H$$

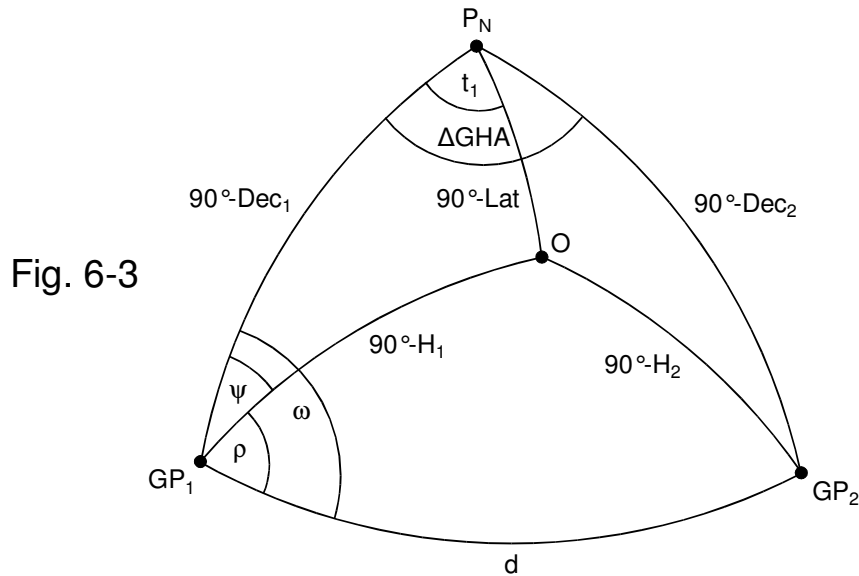
(sun north of observer: +ΔH, sun south of observer: –ΔH)

The exact latitude is obtained by iteration, i. e., we substitute $Lat_{improved}$ for Lat_E and repeat the calculations until the obtained latitude is virtually constant. Usually, no more than one or two iterations are necessary. The method has a few limitations and requires critical judgement. The meridian angle should be small compared with the zenith distance of the sun. Otherwise, a greater number of iterations may be necessary. The method may yield erratic results if Lat_E is similar to Dec. A sight should be discarded when the observer is not sure if the sun is north or south of his position.

The influence of a longitude error on the latitude thus obtained is **not** decreased by iteration.

Latitude by two altitudes

Even if no estimated longitude is available, the exact latitude can still be found by observation of two celestial bodies. The required quantities are Greenwich hour angle, declination, and observed altitude of each body [7].

The calculations are based upon spherical triangles (see chapter 10 & chapter 11). In *Fig. 6-3*, P_N denotes the north pole, O the observer's unknown position, GP_1 the geographic position of the first body, and GP_2 the position of the second body.

First, we consider the spherical triangle [GP_1, P_N, GP_2]. *Fig. 6-3* shows only one of several possible configurations. O may as well be outside the triangle [GP_1, P_N, GP_2]. We form the difference of both Greenwich hour angles, ΔGHA:

$$\Delta GHA = \left|GHA_2 - GHA_1\right|$$

Using the law of cosines for sides (chapter 10), we calculate the great circle distance between GP_1 and GP_2, d.

$$\cos d = \sin Dec_1 \cdot \sin Dec_2 + \cos Dec_1 \cdot \cos Dec_2 \cdot \cos(\Delta GHA)$$

$$d = \arccos\left[\sin Dec_1 \cdot \sin Dec_2 + \cos Dec_1 \cdot \cos Dec_2 \cdot \cos(\Delta GHA)\right]$$

Now we solve the same triangle for the angle ω, the horizontal distance between P_N and GP_2, measured at GP_1:

$$\cos \omega = \frac{\sin Dec_2 - \sin Dec_1 \cdot \cos d}{\cos Dec_1 \cdot \sin d}$$

$$\omega = \arccos\left(\frac{\sin Dec_2 - \sin Dec_1 \cdot \cos d}{\cos Dec_1 \cdot \sin d}\right)$$

For the spherical triangle [GP_1, O, GP_2], we calculate the angle ρ, the horizontal distance between O and GP_2, measured at GP_1.

$$\cos \rho = \frac{\sin H_2 - \sin H_1 \cdot \cos d}{\cos H_1 \cdot \sin d}$$

$$\rho = \arccos\left(\frac{\sin H_2 - \sin H_1 \cdot \cos d}{\cos H_1 \cdot \sin d}\right)$$

We calculate the angle ψ, the horizontal distance between P_N and O, measured at GP_1. There are two solutions (ψ_1 and ψ_2) since cos ρ = cos (-ρ):

$$\psi_1 = |\omega - \rho| \qquad \psi_2 = \omega + \rho$$

The circles of equal altitude intersect each other at two points. The corresponding positions are on opposite sides of the great circle going through GP_1 and GP_2 (not shown in *Fig. 6-3*). Using the law of cosines for sides again, we solve the spherical triangle [GP_1, P_N, O] for Lat. Since we have two solutions for ψ, we obtain two possible latitudes, Lat_1 and Lat_2.

$$\sin Lat_1 = \sin H_1 \cdot \sin Dec_1 + \cos H_1 \cdot \cos Dec_1 \cdot \cos \psi_1$$

$$Lat_1 = \arcsin\left(\sin H_1 \cdot \sin Dec_1 + \cos H_1 \cdot \cos Dec_1 \cdot \cos \psi_1\right)$$

$$\sin Lat_2 = \sin H_1 \cdot \sin Dec_1 + \cos H_1 \cdot \cos Dec_1 \cdot \cos \psi_2$$

$$Lat_2 = \arcsin\left(\sin H_1 \cdot \sin Dec_1 + \cos H_1 \cdot \cos Dec_1 \cdot \cos \psi_2\right)$$

We choose the value nearest to our estimated latitude. The other one is discarded. If both solutions are very similar and a clear distinction is not possible, one of the sights should be discarded, and a body with a more favorable position should be chosen.

Although the method requires more complicated calculations than, e. g., a latitude by Polaris, it has the advantage that measuring two altitudes usually takes less time than finding the maximum altitude of a single body. Moreover, if fixed stars are observed, even a chronometer error of several hours has no significant influence on the resulting latitude since ΔGHA and both declinations change very slowly in this case.

When the horizontal distance between the observed bodies is in the vicinity of 0° or 180°, the observer's position is close to the great circle going through $GP_1$and GP_2. In this case, the two solutions for latitude are similar, and finding which one corresponds with the actual latitude may be difficult (depending on the quality of the estimate). The resulting latitudes are also close to each other when the observed bodies have approximately the same Greenwich hour angle.

Noon Longitude (Longitude by Equal Altitudes)

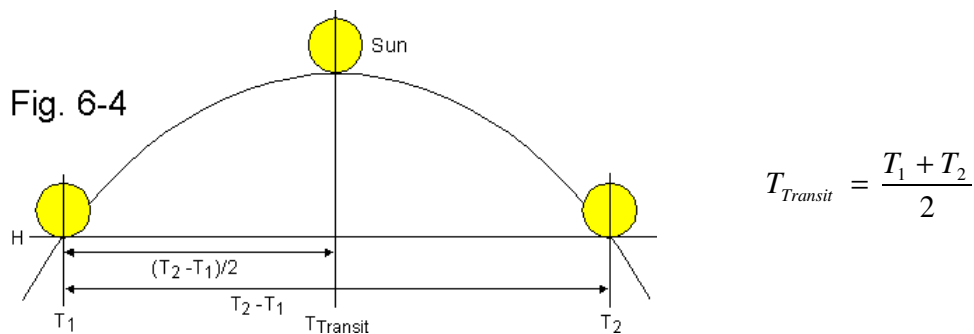
Since the earth rotates with an angular velocity of 15° per hour with respect to the mean sun, the time of local meridian transit (local apparent noon) of the sun, $T_{Transit}$, can be used to calculate the observer's longitude:

$$Lon[°] = 15 \cdot \left(12 - T_{Transit}[h] - EoT_{Transit}[h]\right)$$

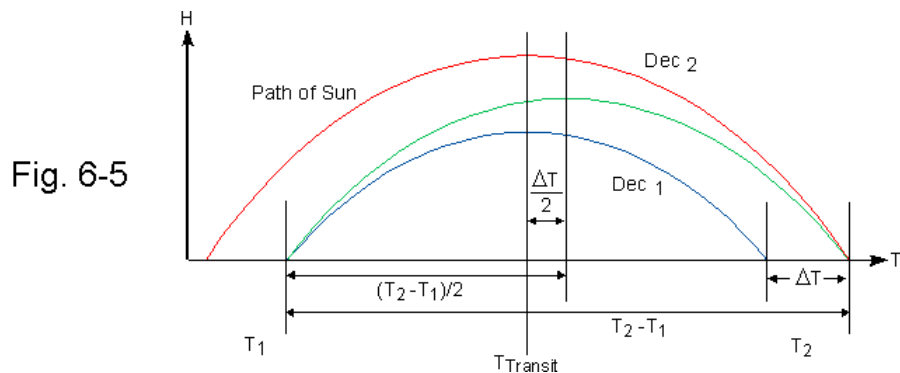
$T_{Transit}$ is measured as GMT (decimal format). The correction for EoT at the time of meridian transit, $EoT_{Transit}$, has to be made because the apparent sun, not the mean sun, is observed (see chapter 3). Since the Nautical Almanac contains only values for EoT (see chapter 3) at 0:00 GMT and 12:00 GMT of each day, $EoT_{Transit}$ has to be found by interpolation.

Since the altitude of the sun - like the altitude of any celestial body - passes through a rather flat maximum, the time of peak altitude is difficult to measure. The exact time of meridian transit can be derived, however, from the times of two equal altitudes of the sun.

Assuming that the sun moves along a symmetrical arc in the sky, $T_{Transit}$ is the mean of the times corresponding with a chosen pair of equal altitudes of the sun, one occurring before LAN (T_1), the other past LAN (T_2) (*Fig. 6-4*):

$$T_{Transit} = \frac{T_1 + T_2}{2}$$

In practice, the times of equal altitudes of the sun are measured as follows:

In the morning, the observer records the time T_1 corresponding with a chosen altitude, H. In the afternoon, the time T_2 is recorded when the descending sun passes through the same altitude again. Since only times of equal altitudes are measured, no altitude correction is required. The interval T_2-T_1 should be greater than approx. 2 hours.

Unfortunately, the arc of the sun is only symmetrical with respect to $T_{Transit}$ if the sun's declination is constant during the observation interval. This is approximately the case around the times of the solstices. During the rest of the year, particularly at the times of the equinoxes, $T_{Transit}$ differs significantly from the mean of T_1 and T_2 due to the changing declination of the sun. *Fig. 6-5* shows the altitude of the sun as a function of time and illustrates how the changing declination affects the apparent path of the sun in the sky, resulting in a time difference, ΔT.

The blue line shows the path of the sun for a given, constant declination, Dec_1. The red line shows how the path would look with a different declination, Dec_2. In both cases, the apparent path of the sun is symmetrical with respect to $T_{Transit}$. However, if the sun's declination varies from Dec_1 at T_1 to Dec_2 at T_2, the path shown by the green line will result.

Now, T_1 and T_2 are no longer symmetrical to $T_{Transit}$. The sun's meridian transit occurs before $(T_1+T_2)/2$ if the sun's declination changes toward the observer's parallel of latitude, like shown in *Fig. 6-5*. Otherwise, the meridian transit occurs after $(T_1+T_2)/2$. Since time and local hour angle (or meridian angle) are proportional to each other, a systematic error in longitude results.

The error in longitude is negligible around the times of the solstices when Dec is almost constant, and is greatest (up to several arcminutes) at the times of the equinoxes when the rate of change of Dec is greatest (approx. 1'/h). Moreover, the error in longitude increases with the observer's latitude and may be quite dramatic in polar regions.

The obtained longitude can be improved, if necessary, by application of the **equation of equal altitudes** [5]:

$$\Delta t \approx \left(\frac{\tan Lat}{\sin t_2} - \frac{\tan Dec_2}{\tan t_2} \right) \cdot \Delta Dec \qquad \Delta Dec = Dec_2 - Dec_1$$

Δt is the change in the meridian angle, t, which cancels the change in altitude resulting from a small change in declination, ΔDec. Lat is the observer's latitude. If the accurate latitude is not known, an estimated latitude may be used. t_2 is the meridian angle of the sun at T_2. Since we do not know the exact value for t_2 initially, we start our calculations with an approximate value calculated from T_1 and T_2:

$$t_2[°] \approx \frac{15 \cdot (T_2[h] - T_1[h])}{2}$$

We denote the improved value for T_2 by T_2^*.

$$T_2^*[h] = T_2[h] - \Delta T[h] = T_2[h] - \frac{\Delta t[°]}{15}$$

At T_2^*, the sun would pass through the same altitude as measured at T_1 if Dec did not change during the interval of observation. Accordingly, the improved time of meridian transit is:

$$T_{Transit} = \frac{T_1 + T_2^*}{2}$$

The residual error resulting from the initial error of t_2 is usually not significant. It can be decreased, if necessary, by iteration. Substituting T_2^* for T_2, we get the improved meridian angle, t_2^*:

$$t_2^*[°] \approx \frac{15 \cdot (T_2^*[h] - T_1[h])}{2}$$

With the improved meridian angle t_2^*, we calculate the improved correction Δt^*:

$$\Delta t^* \approx \left(\frac{\tan Lat}{\sin t_2^*} - \frac{\tan Dec_2}{\tan t_2^*} \right) \cdot \Delta Dec$$

Finally, we obtain a more accurate time value, T_2^{**}:

$$T_2^{**}[h] = T_2[h] - \frac{\Delta t^*[°]}{15}$$

And, accordingly:

$$T_{Transit} = \frac{T_1 + T_2^{**}}{2}$$

The error of ΔDec should be as small as possible. Calculating ΔDec with a high-precision computer almanac is preferable to extracting it from the Nautical Almanac. When using the Nautical Almanac, ΔDec should be calculated from the daily change of declination to keep the rounding error as small as possible.

Although the equation of equal altitudes is strictly valid only for an infinitesimal change of Dec, dDec, it can be used for a measurable change, ΔDec, (up to several arcminutes) as well without sacrificing much accuracy. Accurate time measurement provided, the residual error in longitude rarely exceeds ±0.1'.

Theory of the equation of equal altitudes

The equation of equal altitudes is derived from the altitude formula (see chapter 4) using **differential calculus**:

$$\sin H = \sin Lat \cdot \sin Dec + \cos Lat \cdot \cos Dec \cdot \cos t$$

First, we need to know how a small change in declination would affect sin H. We form the partial derivative with respect to Dec:

$$\frac{\partial(\sin H)}{\partial Dec} = \sin Lat \cdot \cos Dec - \cos Lat \cdot \sin Dec \cdot \cos t$$

Thus, the change in sin H caused by an infinitesimal change in declination, d Dec, is:

$$\frac{\partial(\sin H)}{\partial Dec} \cdot d\,Dec = \left(\sin Lat \cdot \cos Dec - \cos Lat \cdot \sin Dec \cdot \cos t\right) \cdot d\,Dec$$

Now, we form the partial derivative with respect to t in order to find out how a small change in the meridian angle would affect sin H:

$$\frac{\partial(\sin H)}{\partial t} = -\cos Lat \cdot \cos Dec \cdot \sin t$$

The change in sin H caused by an infinitesimal change in the meridian angle, dt, is:

$$\frac{\partial(\sin H)}{\partial t} \cdot d\,t = -\cos Lat \cdot \cos Dec \cdot \sin t \cdot d\,t$$

Since we want both effects to cancel each other, the total differential has to be zero:

$$\frac{\partial(\sin H)}{\partial Dec} \cdot d\,Dec + \frac{\partial(\sin H)}{\partial t} \cdot d\,t = 0$$

$$-\frac{\partial(\sin H)}{\partial t} \cdot d\,t = \frac{\partial(\sin H)}{\partial Dec} \cdot d\,Dec$$

$$\cos Lat \cdot \cos Dec \cdot \sin t \cdot d\,t \;=\; \left(\sin Lat \cdot \cos Dec - \cos Lat \cdot \sin Dec \cdot \cos t\right) \cdot d\,Dec$$

$$d\,t \;=\; \frac{\sin Lat \cdot \cos Dec - \cos Lat \cdot \sin Dec \cdot \cos t}{\cos Lat \cdot \cos Dec \cdot \sin t} \cdot d\,Dec$$

$$d\,t \;=\; \left(\frac{\tan Lat}{\sin t} - \frac{\tan Dec}{\tan t}\right) \cdot d\,Dec$$

$$\Delta t \;\approx\; \left(\frac{\tan Lat}{\sin t} - \frac{\tan Dec}{\tan t}\right) \cdot \Delta\,Dec$$

Longitude Measurement on a Moving Vessel

On a moving vessel, we have to take into account not only the influence of varying declination but also the effects of changing latitude and longitude on the altitude of the body during the observation interval. Differentiating sin H (altitude formula) with respect to Lat, we get:

$$\frac{\partial\left(\sin H\right)}{\partial\,Lat} \;=\; \cos\,Lat \cdot \sin Dec - \sin\,Lat \cdot \cos Dec \cdot \cos\,t$$

Again, the total differential is zero because the combined effects of latitude and meridian angle cancel each other with respect to their influence on sin H:

$$\frac{\partial\left(\sin H\right)}{\partial\,Lat} \cdot d\,Lat + \frac{\partial\left(\sin H\right)}{\partial\,Dec} \cdot d\,t \;=\; 0$$

In analogy with a change in declination, we obtain the following formula for a small change in latitude:

$$d\,t \;=\; \left(\frac{\tan Dec}{\sin t} - \frac{\tan Lat}{\tan t}\right) \cdot d\,Lat$$

The correction for the combined variations in Dec, Lat, and Lon is:

$$\Delta t \;\approx\; \left(\frac{\tan Lat_2}{\sin t_2} - \frac{\tan Dec_2}{\tan t_2}\right) \cdot \Delta Dec \;+\; \left(\frac{\tan Dec_2}{\sin t_2} - \frac{\tan Lat_2}{\tan t_2}\right) \cdot \Delta\,Lat \;-\; \Delta\,Lon$$

ΔLat and ΔLon are the small changes in latitude and longitude corresponding with the path of the vessel traveled between T_1 and T_2. The meridian angle, t_2, has to include a correction for ΔLon:

$$t_2\left[\circ\right] \;\approx\; \frac{15 \cdot \left(T_2\left[h\right] - T_1\left[h\right]\right) - \Delta\,Lon\left[\circ\right]}{2}$$

ΔLat and ΔLon are calculated from the course, C, the velocity over ground, v, and the time elapsed.

$$\Delta\,Lat\left[\,'\,\right] = v\left[\,kn\,\right] \cdot \cos C \cdot \left(T_2\left[\,h\,\right] - T_1\left[\,h\,\right]\right)$$

$$Lat_2 = Lat_1 + \Delta Lat$$

$$\Delta Lon['] = v[kn] \cdot \frac{\sin C}{\cos Lat} \cdot \left(T_2[h] - T_1[h]\right)$$

$$Lon_2 = Lon_1 + \Delta Lon$$

$$1\, kn(knot) = 1\, nm/h$$

C is measured clockwise from true north (0°...360°). Again, the corrected time of equal altitude is:

$$T_2^*[h] = T_2[h] - \frac{\Delta t[°]}{15}$$

$$T_{Transit} = \frac{T_1 + T_2^*}{2}$$

The longitude calculated from $T_{Transit}$ refers to the observer's position at T_1. The longitude at T_2 is Lon+ΔLon.

The longitude error caused by a change in latitude can be dramatic and requires the navigator's particular attention, even if the vessel travels at a moderate speed. The above considerations clearly demonstrate that determining one's **exact** longitude by equal altitudes of the sun is not as simple as it seems to be at first glance, particularly on a traveling vessel. It is therefore quite natural that with the development of position line navigation (including simple graphic solutions for a traveling vessel), longitude by equal altitudes became less important.

The Meridian Angle of the Sun at Maximum Altitude

Fig. 6-5 shows that the maximum altitude of the sun is slightly different from the altitude at the moment of meridian passage if the declination changes. At maximum altitude, the rate of change of altitude caused by the changing declination cancels the rate of change of altitude caused by the changing meridian angle.
The equation of equal altitude enables us to calculate the meridian angle of the sun at this moment. We divide each side of the equation by the infinitesimal time interval dT:

$$\frac{dt}{dT} = \left(\frac{\tan Lat}{\sin t} - \frac{\tan Dec}{\tan t}\right) \cdot \frac{d\,Dec}{dT}$$

Measuring the rate of change of t and Dec in arcminutes per hour we get:

$$900\,'/h = \left(\frac{\tan Lat}{\sin t} - \frac{\tan Dec}{\tan t}\right) \cdot \frac{d\,Dec[']}{dT[h]}$$

Since t is a very small angle, we can substitute tan t for sin t:

$$900 \approx \frac{\tan Lat - \tan Dec}{\tan t} \cdot \frac{d\,Dec[']}{dT[h]}$$

Now, we can solve the equation for tan t:

$$\tan t \approx \frac{\tan Lat - \tan Dec}{900} \cdot \frac{d\,Dec[']}{dT[h]}$$

Since a small angle (in radians) is nearly equal to its tangent, we get:

$$t\left[°\right] \cdot \frac{\pi}{180} \approx \frac{\tan Lat - \tan Dec}{900} \cdot \frac{d\,Dec\left['\right]}{d\,T\left[h\right]}$$

Measuring t in arcminutes, the equation is stated as:

$$t\left['\right] \approx 3.82 \cdot \left(\tan Lat - \tan Dec\right) \cdot \frac{d\,Dec\left['\right]}{d\,T\left[h\right]}$$

dDec/dT is the rate of change of declination measured in arcminutes per hour.

The maximum altitude occurs after meridian transit if t is positive, and before meridian transit if t is negative.

For example, at the time of the spring equinox (Dec ≈ 0, dDec/dT ≈ +1'/h) an observer being at +80° (N) latitude would observe the maximum altitude of the sun at t ≈ +21.7', i. e., 86.8 seconds after meridian transit (LAN). An observer at +45° latitude, however, would observe the maximum altitude at t ≈ +3.82', i. e., only 15.3 seconds after meridian transit.

The Maximum Altitude of the Sun

We can use the last equation to evaluate the systematic error of a noon latitude. The latter is based upon the maximum altitude of the sun, not on the altitude at the moment of meridian transit. Following the above example, the observer at 80° latitude would observe the maximum altitude 86.7 seconds after meridian transit. During this interval, the declination of the sun would have changed from 0 to +1.445" (assuming that Dec is 0 at the time of meridian transit). Using the altitude formula (chapter 4), we get:

$$Hc = \arcsin\left(\sin 80° \cdot \sin 1.445'' + \cos 80° \cdot \cos 1.445'' \cdot \cos 21.7'\right) = 10°\ 0'\ 0.72''$$

In contrast, the calculated altitude at meridian transit would be exactly 10°. Thus, the error of the noon latitude would be -0.72".

In the same way, we can calculate the maximum altitude of the sun observed at 45° latitude:

$$Hc = \arcsin\left(\sin 45° \cdot \sin 0.255'' + \cos 45° \cdot \cos 0.255'' \cdot \cos 3.82'\right) = 45°\ 0'\ 0.13''$$

In this case, the error of the noon latitude would be only -0.13".

The above examples show that even at the times of the equinoxes, the systematic error of a noon latitude caused by the changing declination of the sun is not significant because it is much smaller than other observational errors, e. g., the errors in dip or refraction. A measurable error in latitude can only occur if the observer is very close to one of the poles (tan Lat!). Around the times of the solstices, the error in latitude is practically non-existent.

Time Sight

The process of deriving the longitude from a single altitude of a body (as well as the observation made for this purpose) is called **time sight**. However, this method requires knowledge of the exact latitude, e. g., a noon latitude. Solving the navigational triangle (chapter 11) for the meridian angle, t, we get:

$$t = \pm \arccos \frac{\sin Ho - \sin Lat \cdot \sin Dec}{\cos Lat \cdot \cos Dec}$$

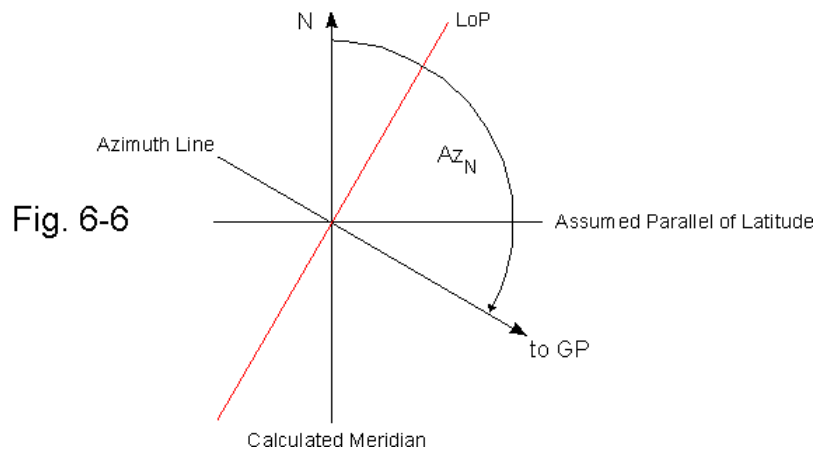
The equation has two solutions, +t and –t, since cos t = cos (–t). Geometrically, this corresponds with the fact that the circle of equal altitude intersects the parallel of latitude at two points.

Using the following formulas and rules, we obtain the longitudes of these points of intersection, Lon_1 and Lon_2:

$$Lon_1 = t - GHA$$

$$Lon_2 = 360° - t - GHA$$

$$\text{If } Lon_1 < -180° \rightarrow Lon_1 + 360°$$
$$\text{If } Lon_2 < -180° \rightarrow Lon_2 + 360°$$
$$\text{If } Lon_2 > +180° \rightarrow Lon_2 - 360°$$

Even if we do not know the exact latitude, we can still use a time sight to derive a line of position from an assumed latitude. After solving the time sight, we plot the assumed parallel of latitude and the calculated meridian.
Next, we calculate the azimuth of the body with respect to the position thus obtained (azimuth formulas, chapter 4) and plot the azimuth line. Our line of position is the perpendicular of the azimuth line going through the calculated position (*Fig. 6-6*).

The latter method is of historical interest only. The modern navigator will certainly prefer the intercept method (chapter 4) which can be used without any restrictions regarding meridian angle (local hour angle), latitude, and declination (see below).

A time sight is not reliable when the body is close to the meridian. Using differential calculus, we can demonstrate that the error of the meridian angle, dt, resulting from an altitude error, dH, varies in proportion with 1/sin t:

$$dt = -\frac{\cos Ho}{\cos Lat \cdot \cos Dec \cdot \sin t} \cdot dH$$

Moreover, dt varies inversely with cos Lat and cos Dec. Therefore, high latitudes and declinations should be avoided as well. The same restrictions apply to *Sumner's* method which is based upon two time sights.

Direct Computation of Position

If we know the exact time, the observations for a **latitude by two altitudes** even enable us to calculate our position directly, without any graphic plot. After obtaining our latitude, Lat, from two altitudes (see above), we use the time sight formula to calculate the meridian angle of one of the bodies. In case of the first body, for example, we calculate t_1 from the quantities Lat, Dec_1, and H_1 (see *Fig. 6-3*). Two possible longitudes result from the meridian angle thus obtained. We choose the one nearest to our estimated longitude. This is a **rigorous method**, not an approximation. Direct computation was rarely used in the past since the calculations are more complicated than those required for graphic solutions. Of course, in the age of computers the complexity of the method does not pose a problem anymore.

Chapter 7

Finding Time and Longitude by Lunar Distances

In celestial navigation, time and longitude are interdependent. Finding one's longitude at sea or in unknown terrain is impossible without knowing the exact time and vice versa. Therefore, old-time navigators were basically restricted to latitude sailing on long voyages, i. e., they had to sail along a chosen parallel of latitude until they came in sight of the coast. Since there was no reliable estimate of the time of arrival, many ships ran ashore during periods of darkness or bad visibility. Spurred by heavy losses of men and material, scientists tried to solve the longitude problem by using astronomical events as time marks. In principle, such a method is only suitable when the observed time of the event is virtually independent of the observer's geographic position.

Measuring time by the apparent movement of the moon with respect to the background of fixed stars was suggested in the 15th century already (*Regiomontanus*) but proved impracticable since neither reliable ephemerides for the moon nor precise instruments for measuring angles were available at that time.

Around the middle of the 18th century, astronomy and instrument making had finally reached a stage of development that made time measurement by lunar observations possible. Particularly, deriving the time from a so-called **lunar distance**, the angular distance of the moon from a chosen reference body, became a popular method. Although the procedure is rather cumbersome, it became an essential part of celestial navigation and was used far into the 19th century, long after the invention of the mechanical chronometer (*Harrison*, 1736). This was mainly due to the limited availability of reliable chronometers and their exorbitant price. When chronometers became affordable around the middle of the 19th century, lunar distances gradually went out of use. Until 1906, the Nautical Almanac included lunar distance tables showing predicted geocentric angular distances between the moon and selected bodies in 3-hour intervals.* After the tables were dropped, lunar distances fell more or less into oblivion. Not much later, radio time signals became available world-wide, and the longitude problem was solved once and for all. Today, lunar distances are mainly of historical interest. The method is so ingenious, however, that a detailed study is worthwhile.

The basic idea of the lunar distance method is easy to comprehend. Since the moon moves across the celestial sphere at a rate of about 0.5° per hour, the angular distance between the moon, M, and a body in her path, B, varies at a similar rate and rapidly enough to be used to measure the time. The time corresponding with an observed lunar distance can be found by comparison with tabulated values.

Tabulated lunar distances are calculated from the geocentric equatorial coordinates of M and B using the cosine law:

$$\cos D = \sin Dec_M \cdot \sin Dec_B + \cos Dec_M \cdot \cos Dec_B \cdot \cos\left(GHA_M - GHA_B\right)$$

or

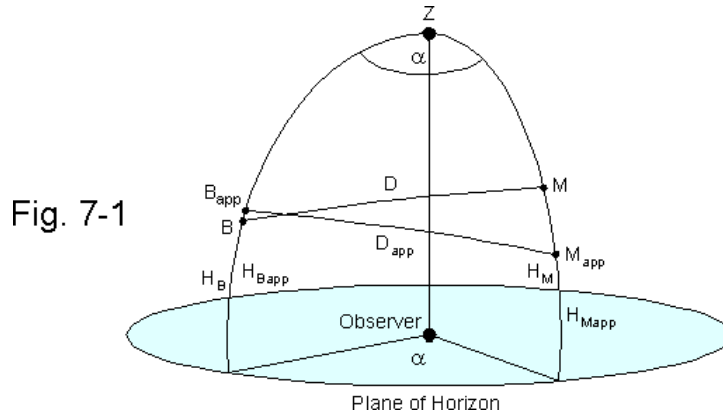
$$\cos D = \sin Dec_M \cdot \sin Dec_B + \cos Dec_M \cdot \cos Dec_B \cdot \cos\left[15 \cdot \left(RA_M[h] - RA_B[h]\right)\right]$$

D is the geocentric lunar distance. These formulas can be used to set up one's own table with the aid of the Nautical Almanac or any computer almanac if a lunar distance table is not available.

*Almost a century after the original Lunar Distance Tables were dropped, Steven Wepster resumed the tradition. His tables are presently (2004) available through the internet [14].

Clearing the lunar distance

Before a lunar distance measured by the observer can be compared with tabulated values, it has to be reduced to the corresponding geocentric angle by clearing it from the effects of refraction and parallax. This essential process is called **clearing the lunar distance**. Numerous procedures have been developed, among them rigorous and "quick" methods. In the following, we will discuss the almost identical methods by *Dunthorne* (1766) and *Young* (1856). They are rigorous for a spherical model of the earth.

Fig. 7-1 shows the positions of the moon and a reference body in the coordinate system of the horizon. We denote the apparent positions of the centers of the moon and the reference body by M_{app} and B_{app}, respectively. Z is the zenith.

The side D_{app} of the spherical triangle B_{app}-Z-M_{app} is the apparent lunar distance. The altitudes of M_{app} and B_{app} (obtained after applying the corrections for index error, dip, and semidiameter) are H_{Mapp} and H_{Bapp}, respectively. The vertical circles of both bodies form the angle α, the difference between the azimuth of the moon, Az_M, and the azimuth of the reference body, Az_B:

$$\alpha = Az_M - Az_B$$

The position of each body is shifted along its vertical circle by atmospheric refraction and parallax in altitude. After correcting H_{Mapp} and H_{Bapp} for both effects, we obtain the geocentric positions M and B. We denote the altitude of M by H_M and the altitude of B by H_B. H_M is always greater than H_{Mapp} because the parallax of the moon is always greater than refraction. The angle α is neither affected by refraction nor by the parallax in altitude:

$$Az_M = Az_{Mapp} \qquad Az_B = Az_{Bapp}$$

The side D of the spherical triangle B-Z-M is the unknown geocentric lunar distance. If we knew the exact value for α, calculation of D would be very simple (cosine law). Unfortunately, the navigator has no means for measuring α precisely. It is possible, however, to calculate D solely from the five quantities D_{app}, H_{Mapp}, H_M, H_{Bapp}, and H_B.

Applying the cosine formula to the spherical triangle formed by the zenith and the apparent positions, we get:

$$\cos D_{app} = \sin H_{Mapp} \cdot \sin H_{Bapp} + \cos H_{Mapp} \cdot \cos H_{Bapp} \cdot \cos \alpha$$

$$\cos \alpha = \frac{\cos D_{app} - \sin H_{Mapp} \cdot \sin H_{Bapp}}{\cos H_{Mapp} \cdot \cos H_{Bapp}}$$

Repeating the procedure with the spherical triangle formed by the zenith and the geocentric positions, we get:

$$\cos D = \sin H_M \cdot \sin H_B + \cos H_M \cdot \cos H_B \cdot \cos \alpha$$

$$\cos \alpha = \frac{\cos D - \sin H_M \cdot \sin H_B}{\cos H_M \cdot \cos H_B}$$

Since α is constant, we can combine both azimuth formulas:

$$\frac{\cos D - \sin H_M \cdot \sin H_B}{\cos H_M \cdot \cos H_B} = \frac{\cos D_{app} - \sin H_{Mapp} \cdot \sin H_{Bapp}}{\cos H_{Mapp} \cdot \cos H_{Bapp}}$$

Thus, we have eliminated the unknown angle α. Now, we subtract unity from both sides of the equation:

$$\frac{\cos D - \sin H_M \cdot \sin H_B}{\cos H_M \cdot \cos H_B} - 1 = \frac{\cos D_{app} - \sin H_{Mapp} \cdot \sin H_{Bapp}}{\cos H_{Mapp} \cdot \cos H_{Bapp}} - 1$$

$$\frac{\cos D - \sin H_M \cdot \sin H_B}{\cos H_M \cdot \cos H_B} - \frac{\cos H_M \cdot \cos H_B}{\cos H_M \cdot \cos H_B} = \frac{\cos D_{app} - \sin H_{Mapp} \cdot \sin H_{Bapp}}{\cos H_{Mapp} \cdot \cos H_{Bapp}} - \frac{\cos H_{Mapp} \cdot \cos H_{Bapp}}{\cos H_{Mapp} \cdot \cos H_{Bapp}}$$

$$\frac{\cos D - \sin H_M \cdot \sin H_B - \cos H_M \cdot \cos H_B}{\cos H_M \cdot \cos H_B} = \frac{\cos D_{app} - \sin H_{Mapp} \cdot \sin H_{Bapp} - \cos H_{Mapp} \cdot \cos H_{Bapp}}{\cos H_{Mapp} \cdot \cos H_{Bapp}}$$

Using the addition formula for cosines, we have:

$$\frac{\cos D - \cos\left(H_M - H_B\right)}{\cos H_M \cdot \cos H_B} = \frac{\cos D_{app} - \cos\left(H_{Mapp} - H_{Bapp}\right)}{\cos H_{Mapp} \cdot \cos H_{Bapp}}$$

Solving for cos D, we obtain *Dunthorne's* formula for clearing the lunar distance:

$$\cos D = \frac{\cos H_M \cdot \cos H_B}{\cos H_{Mapp} \cdot \cos H_{Bapp}} \cdot \left[\cos D_{app} - \cos\left(H_{Mapp} - H_{Bapp}\right)\right] + \cos\left(H_M - H_B\right)$$

Adding unity to both sides of the equation instead of subtracting it, leads to *Young's* formula:

$$\cos D = \frac{\cos H_M \cdot \cos H_B}{\cos H_{Mapp} \cdot \cos H_{Bapp}} \cdot \left[\cos D_{app} + \cos\left(H_{Mapp} + H_{Bapp}\right)\right] - \cos\left(H_M + H_B\right)$$

Procedure

Deriving UT from a lunar distance comprises the following steps:

1.

We measure the altitude of the upper or lower limb of the moon, whichever is visible, and note the watch time of the observation, $WT1_{LMapp}$.
We apply the corrections for index error and dip (if necessary) and get the apparent altitude of the limb, $H1_{LMapp}$. We repeat the procedure with the reference body and obtain the watch time $WT1_{Bapp}$ and the altitude $H1_{Bapp}$.

2.

We measure the angular distance between the limb of the moon and the reference body, D_{Lapp}, and note the corresponding watch time, WT_D. The angle D_{Lapp} has to be measured with the greatest possible precision. It is recommended to measure a few D_{Lapp} values and their corresponding WT_D values in rapid succession and calculate the respective average value. When the moon is almost full, it is not quite easy to distinguish the limb of the moon from the terminator (shadow line). In general, the limb has a sharp appearance whereas the terminator is slightly indistinct.

3.

We measure the altitudes of both bodies again, as described above. We denote them by $H2_{LMapp}$ and $H2_{Bapp}$, and note the corresponding watch times of observation, $WT2_{LMapp}$ and $WT2_{Bapp}$.

4.

Since the observations are only a few minutes apart, we can calculate the altitude of the respective body at the moment of the lunar distance observation by linear interpolation:

$$H_{LMapp} = H1_{LMapp} + \left(H2_{LMapp} - H1_{LMapp}\right) \cdot \frac{WT_D - WT1_{LMapp}}{WT2_{LMapp} - WT1_{LMapp}}$$

$$H_{Bapp} = H1_{Bapp} + \left(H2_{Bapp} - H1_{Bapp}\right) \cdot \frac{WT_D - WT1_{Bapp}}{WT2_{Bapp} - WT1_{Bapp}}$$

5.

We correct the altitude of the moon and the angular distance D_{Lapp} for the augmented semidiameter of the moon, SD_{aug}. The latter can be calculated directly from the altitude of the upper or lower limb of the moon:

$$\tan SD_{aug} = \frac{k}{\sqrt{\frac{1}{\sin^2 HP_M} - \left(\cos H_{LMapp} \pm k\right)^2} - \sin H_{LMapp}} \qquad k = 0.2725$$

$$(\text{upper limb: } \cos H_{LMapp} - k \qquad \text{lower limb: } \cos H_{LMapp} + k)$$

The altitude correction is:

$$\text{Lower limb: } \quad H_{Mapp} = H_{LMapp} + SD_{aug}$$

$$\text{Upper limb: } \quad H_{Mapp} = H_{LMapp} - SD_{aug}$$

The rules for the lunar distance correction are:

$$\text{Limb of moon towards reference body: } \quad D_{app} = D_{Lapp} + SD_{aug}$$

$$\text{Limb of moon away from reference body: } \quad D_{app} = D_{Lapp} - SD_{aug}$$

The above procedure is an approximation since the augmented semidiameter is a function of the altitude corrected for refraction. Since refraction is a small quantity and since the total augmentation between 0° and 90° altitude is only approx. 0.3', the resulting error is very small and may be ignored.

The sun, when chosen as reference body, requires the same corrections for semidiameter. Since the sun does not show a measurable augmentation, we can use the geocentric semidiameter tabulated in the Nautical Almanac or calculated with a computer program.

6.

We correct both altitudes, H_{Mapp} and H_{Bapp}, for atmospheric refraction, R.

$$R_i\,['] = \frac{p\,[mbar]}{1010}\cdot\frac{283}{T\,[°C]+273}\cdot\left(\frac{0.97127}{\tan H_i}-\frac{0.00137}{\tan^3 H_i}\right) \qquad i = Mapp,\, Bapp \qquad H_i > 10°$$

R_i is subtracted from the respective altitude. The refraction formula is only accurate for altitudes above approx. 10°. Lower altitudes should be avoided anyway since refraction may become erratic and since the apparent disk of the moon (and sun) assumes an oval shape caused by an increasing difference in refraction for upper and lower limb. This distortion would affect the semidiameter with respect to the reference body in a complicated way.

7.

We correct the altitudes for the parallax in altitude:

$$\sin P_M = \sin HP_M \cdot \cos\left(H_{Mapp} - R_{Mapp}\right) \qquad \sin P_B = \sin HP_B \cdot \cos\left(H_{Bapp} - R_{Bapp}\right)$$

We apply the altitude corrections as follows:

$$H_M = H_{Mapp} - R_{Mapp} + P_M \qquad H_B = H_{Bapp} - R_{Bapp} + P_B$$

The correction for parallax is not applied to the altitude of a fixed star ($HP_B = 0$).

8.

With D_{app}, H_{Mapp}, H_M, H_{Bapp}, and H_B, we calculate D using *Dunthorne's* or *Young's* formula.

9.

The time corresponding with the geocentric distance D is found by interpolation. Lunar distance tables show D as a function of time, T (UT). If the rate of change of D does not vary too much (less than approx. 0.3' in 3 hours), we can use linear interpolation. However, in order to find T, we have to consider T as a function of D (inverse interpolation).

$$T_D = T_1 + \left(T_2 - T_1\right)\cdot\frac{D - D_1}{D_2 - D_1}$$

T_D is the unknown time corresponding with D. D_1 and D_2 are tabulated lunar distances. T_1 and T_2 are the corresponding time (UT) values ($T_2 = T_1 + 3h$). D is the geocentric lunar distance calculated from D_{app}. D has to be between D_1 and D_2.

If the rate of change of D varies significantly, more accurate results are obtained with methods for non-linear interpolation, for example, with 3-point *Lagrange* interpolation. Choosing three pairs of tabulated values, (T_1, D_1), (T_2, D_2), and (T_3, D_3), T_D is calculated as follows:

$$T_D = T_1\cdot\frac{\left(D - D_2\right)\cdot\left(D - D_3\right)}{\left(D_1 - D_2\right)\cdot\left(D_1 - D_3\right)} + T_2\cdot\frac{\left(D - D_1\right)\cdot\left(D - D_3\right)}{\left(D_2 - D_1\right)\cdot\left(D_2 - D_3\right)} + T_3\cdot\frac{\left(D - D_1\right)\cdot\left(D - D_2\right)}{\left(D_3 - D_1\right)\cdot\left(D_3 - D_2\right)}$$

$$T_2 = T_1 + 3h, \qquad T_3 = T_2 + 3h, \qquad D_1 < D_2 < D_3 \;\; or \;\; D_1 > D_2 > D_3$$

D may have any value between D_1 and D_3.

There must not be a minimum or maximum of D in the time interval [T_1, T_3]. This problem does not occur with a properly chosen body having a suitable rate of change of D. Near a minimum or maximum of D, ΔD/ΔT would be very small, and the observation would be erratic anyway.
After finding T_D, we can calculate the watch error, ΔT.

$$\Delta T = WT_D - T_D$$

ΔT is the difference between our watch time at the moment of observation, WT_D, and the time found by interpolation, T_D.

Subtracting the watch error from the watch time, WT, results in UT.

$$UT = WT - \Delta T$$

Improvements

The procedures described so far refer to a spherical earth. In reality, however, the earth has approximately the shape of an ellipsoid flattened at the poles. This leads to small but measurable effects when observing the moon, the body nearest to the earth. First, the parallax in altitude differs slightly from the value calculated for a spherical earth. Second, there is a small parallax in azimuth which would not exist if the earth were a sphere (see chapter 9). If no correction is applied, D may contain an error of up to approx. 0.2'. The following formulas refer to an observer on the surface of the reference ellipsoid (approximately at sea level).

The corrections require knowledge of the observer's latitude, Lat, the true azimuth of the moon, Az_M, and the true azimuth of the reference body, Az_B.

Since the corrections are small, the three values do not need to be very accurate. Errors of a few degrees are tolerable. Instead of the azimuth, the compass bearing of each body, corrected for magnetic declination, may be used.

Parallax in altitude:

This correction is applied to the parallax in altitude and is used to calculate H_M with higher precision before clearing the lunar distance.

$$\Delta P_M \approx f \cdot HP_M \cdot \left[\sin(2 \cdot Lat) \cdot \cos Az_M \cdot \sin H_{Mapp} - \sin^2 Lat \cdot \cos H_{Mapp} \right]$$

f is the flattening of the earth: $f = \frac{1}{298.257}$

$$P_{M,improved} = P_M + \Delta P_M$$

$$H_M = H_{Mapp} - R_{Mapp} + P_{M,improved}$$

Parallax in azimuth:

The correction for the parallax in azimuth is applied after calculating H_M and D. The following formula is a fairly accurate approximation of the parallax in azimuth, ΔAz_M:

$$\Delta Az_M \approx f \cdot HP_M \cdot \frac{\sin(2 \cdot Lat) \cdot \sin Az_M}{\cos H_M}$$

In order to find, how ΔAz_M affects D, we go back to the cosine formula:

$$\cos D = \sin H_M \cdot \sin H_B + \cos H_M \cdot \cos H_B \cdot \cos \alpha$$

We differentiate the equation with respect to α:

$$\frac{d(\cos D)}{d\alpha} = -\cos H_M \cdot \cos H_B \cdot \sin \alpha$$

$$d(\cos D) = -\sin D \cdot dD$$

$$-\sin D \cdot dD = -\cos H_M \cdot \cos H_B \cdot \sin \alpha \cdot d\alpha$$

$$dD = \frac{\cos H_M \cdot \cos H_B \cdot \sin \alpha}{\sin D} \cdot d\alpha$$

Since $d\alpha = d\,Az_M$, the change in D caused by an infinitesimal change in Az_M is:

$$dD = \frac{\cos H_M \cdot \cos H_B \cdot \sin \alpha}{\sin D} \cdot dAz_M$$

With a small but measurable change in Az_M, we have:

$$\Delta D \approx \frac{\cos H_M \cdot \cos H_B \cdot \sin \alpha}{\sin D} \cdot \Delta Az_M$$

$$D_{improved} \approx D + \Delta D$$

Combining the formulas for ΔAz_M and ΔD, we get:

$$D_{improved} \approx D + f \cdot HP_M \cdot \frac{\cos H_B \cdot \sin(2 \cdot Lat) \cdot \sin Az_M \cdot \sin(Az_M - Az_B)}{\sin D}$$

Accuracy

According to modern requirements, the lunar distance method is horribly inaccurate. In the 18th and early 19th century, however, this was generally accepted because a longitude with an error of 0.5°-1° was still better than no longitude at all. Said error is the approximate result of an error of only 1' in the measurement of D_{Lapp}, not uncommon for a sextant reading under practical conditions. Therefore, D_{Lapp} should be measured with greatest care.

The altitudes of both bodies do not quite require the same degree of precision because a small error in the apparent altitude leads to about the same error in the geocentric altitude. Since both errors cancel each other to a large extent, the resulting error in D is comparatively small. An altitude error of a few arcminutes is tolerable in most cases. Therefore, measuring two altitudes of each body and finding the altitude at the moment of the lunar distance observation by interpolation is not absolutely necessary. Measuring a single altitude of each body shortly before or after the lunar distance measurement is sufficient if a small loss in accuracy is accepted.

The position of the reference body with respect to the moon is crucial. The standard deviation of a time value obtained by lunar distance is inversely proportional to the rate of change of D. Since the plane of the lunar orbit forms a relatively small angle (approx. 5°) with the ecliptic, bright bodies in the vicinity of the ecliptic are most suitable (sun, planets, selected stars).

The stars generally recommended for the lunar distance method are Aldebaran, Altair, Antares, Fomalhaut, Hamal, Markab, Pollux, Regulus, and Spica, but other stars close to the ecliptic may be used as well, e. g., Nunki. The lunar distance tables of the *Nautical Almanac* contained only D values for those bodies having a favorable position with respect to the moon on the day of observation. If in doubt, the navigator should check the rate of change of D. The latter becomes zero when D passes through a minimum or maximum, making an observation useless.

Chapter 8

Rise, Set, Twilight

General Visibility

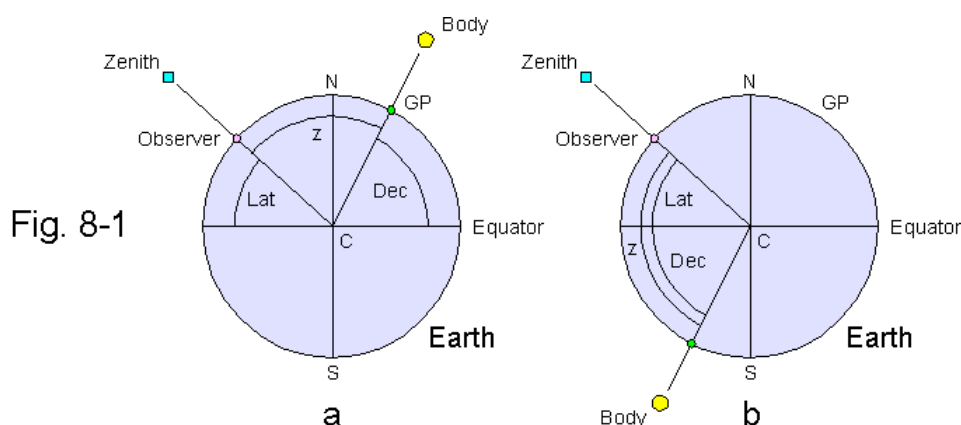
For the planning of observations, it is useful to know the times during which a certain body is above the horizon as well as the times of sunrise, sunset, and twilight.

A body can be always above the horizon, always below the horizon, or above the horizon during a part of the day, depending on the observer's latitude and the declination of the body.

A body is **circumpolar** (always above the celestial horizon) if the zenith distance is smaller than 90° at the moment of **lower** meridian passage, i. e., when the body is on the lower branch of the local meridian (*Fig 8-1a*). This is the case if

$$\left| Lat + Dec \right| > 90°$$

A body is continually below the celestial horizon if the zenith distance is greater than 90° at the instant of **upper** meridian passage (*Fig 8-1b*). This is the case if

$$\left| Lat - Dec \right| > 90°$$

A celestial body being on the same hemisphere as the observer is either sometimes above the horizon or circumpolar. A body being on the opposite hemisphere is either sometimes above the horizon or permanently invisible, but never circumpolar.

The sun provides a good example of how the visibility of a body is affected by latitude and declination. At the time of the summer solstice (Dec = +23.5°), the sun is circumpolar to an observer being north of the **arctic circle** (Lat > +66.5°). At the same time, the sun remains below the celestial horizon all day if the observer is south of the **antarctic circle** (Lat < −66.5°). At the times of the equinoxes (Dec = 0°), the sun is circumpolar only at the poles. At the time of the winter solstice (Dec = −23.5°), the sun is circumpolar south of the antarctic circle and invisible north of the arctic circle. If the observer is between the arctic and the antarctic circle, the sun is visible during a part of the day all year round.

Rise and Set

The events of rise and set can be used to determine latitude, longitude, or time. One should not expect very accurate results, however, since the atmospheric refraction may be erratic if the body is on or near the horizon.

The geometric rise or set of a body occurs when the center of the body passes through the celestial horizon (H = 0°). Due to the influence of atmospheric refraction, all bodies except the moon appear above the visible and sensible horizon at this instant. The moon is not visible at the moment of her geometric rise or set since the depressing effect of the horizontal parallax (~1°) is greater than the elevating effect of atmospheric refraction.

The approximate apparent altitudes (referring to the sensible horizon) at the moment of the astronomical rise or set are:

Sun (lower limb): 15'

Stars: 29'

Planets: 29' – HP

When measuring these altitudes with reference to the sea horizon, we have to add the dip of horizon (chapter 2) to the above values. For example, the altitude of the lower limb of the rising or setting sun is approx. 20' if the height of eye is 8m.

We begin with the well-known altitude formula (see chapter 4).

$$\sin H \;=\; 0 \;=\; \sin Lat \cdot \sin Dec + \cos Lat \cdot \cos Dec \cdot \cos\; t$$

$$\cos\; t \;=\; -\frac{\sin Lat \cdot \sin Dec}{\cos Lat \cdot \cos Dec}$$

Solving the equation for the meridian angle, t, we get :

$$t \;=\; \arccos\; \left(-\tan Lat \cdot \tan Dec \right)$$

The equation has no solution if the argument of the inverse cosine is smaller than −1 or greater than 1. In the first case, the body is **circumpolar**, in the latter case, the body remains continuously below the horizon. Otherwise, the arccos function returns values in the range from 0° through 180°.

Due to the ambiguity of the arccos function, the equation has two solutions, one for rise and one for set. For the calculations below, we have to observe the following rules:

If the body is **rising** (body eastward from the observer), t is treated as a **negative** quantity.

If the body is **setting** (body westward from the observer), t is treated as a **positive** quantity.

If we know our latitude and the time of rise or set, we can calculate our longitude:

$$Lon \;=\; \pm t - GHA$$

GHA is the Greenwich hour angle of the body at the moment of rise or set. The sign of t has to be observed carefully (see above). If the resulting longitude is smaller than −180°, we add 360°.

Knowing our position, we can calculate the times of sunrise and sunset:

$$GMT_{Surise/set} \;=\; 12 \;\pm\; \frac{t\left[\,°\,\right]}{15} - \frac{Lon\left[\,°\,\right]}{15} - EoT$$

The times of sunrise and sunset obtained with the above formula are not quite accurate since Dec and EoT are variable. Since we do not know the exact time of rise or set at the beginning, we have to use estimated values for Dec and EoT initially. The time of rise or set can be improved by iteration (repeating the calculations with Dec and EoT at the calculated time of rise or set). Further, the times thus calculated are influenced by the irregularities of atmospheric refraction near the horizon. Therefore, a time error of ±2 minutes is not unusual.

Accordingly, we can calculate our longitude from the time of sunrise or sunset if we know our latitude:

$$Lon[°] = \pm t + 15 \cdot (12 - GMT_{Sunrise/set} - EoT)$$

Again, this is not a very precise method, and an error of several arcminutes in longitude is not unlikely.

Knowing our longitude, we are able to determine our approximate latitude from the time of sunrise or sunset:

$$t[°] = Lon[°] - 15 \cdot (12 - GMT_{Sunrise/set} - EoT)$$

$$Lat = \arctan\left(-\frac{\cos t}{\tan Dec}\right)$$

In navigation, rise and set are defined as the moments when the upper limb of a body is on the visible horizon. These events can be observed without a sextant. Now, we have to take into account the effects of refraction, horizontal parallax, dip, and semidiameter. These quantities determine the altitude (Ho) of a body with respect to the celestial horizon at the instant of the visible rise or set.

$$t = \arccos\frac{\sin Ho - \sin Lat \cdot \sin Dec}{\cos Lat \cdot \cos Dec}$$

$$Ho = HP - SD - R_H - Dip$$

According to the Nautical Almanac, the refraction for a body being on the sensible horizon, R_H, is approximately (!) 34'.

When observing the upper limb of the sun, we get:

$$Ho = 0.15' - 16' - 34' - Dip \approx -50' - Dip$$

Ho is negative. If we refer to the upper limb of the sun and the sensible horizon (Dip=0), the meridian angle at the time of sunrise or sunset is:

$$t = \arccos\frac{-0.0145 - \sin Lat \cdot \sin Dec}{\cos Lat \cdot \cos Dec}$$

Azimuth and Amplitude

The azimuth angle of a rising or setting body is calculated with the azimuth formula (see chapter 4):

$$Az = \arccos\frac{\sin Dec - \sin H \cdot \sin Lat}{\cos H \cdot \cos Lat}$$

With H=0, we get:

$$Az = \arccos \frac{\sin Dec}{\cos Lat}$$

Az is +90° (rise) and –90° (set) if the declination of the body is zero, regardless of the observer's latitude. Accordingly, the sun rises in the east and sets in the west at the times of the equinoxes (geometric rise and set).

With H_{center}= –50' (upper limb of the sun on the sensible horizon), we have:

$$Az = \arccos \frac{\sin Dec + 0.0145 \cdot \sin Lat}{0.9999 \cdot \cos Lat}$$

The true azimuth of the rising or setting body is:

$$Az_N = \begin{cases} Az & \text{if} \quad t < 0 \\ 360° - Az & \text{if} \quad t > 0 \end{cases}$$

The azimuth of a body at the moment of rise or set can be used to find the magnetic declination at the observer's position (compare with chapter 13).

The horizontal angular distance of the rising/setting body from the east/west point on the horizon is called **amplitude** and can be calculated from the azimuth. An amplitude of E45°N, for instance, means that the body rises 45° north of the east point on the horizon.

Twilight

At sea, twilight is important for the observation of stars and planets since it is the only time when these bodies **and** the horizon are visible. By definition, there are three kinds of twilight. The altitude, H, refers to the center of the sun and the celestial horizon and marks the beginning (morning) and the end (evening) of the respective twilight.

Civil twilight: H = –6°

Nautical twilight: H = –12°

Astronomical twilight: H = –18°

In general, an altitude of the sun between –3° and –9° is recommended for astronomical observations at sea (best visibility of brighter stars and sea horizon). However, exceptions to this rule are possible, depending on the actual weather conditions.

The meridian angle for the sun at –6° altitude (center) is:

$$t = \arccos \frac{-0.10453 - \sin Lat \cdot \sin Dec}{\cos Lat \cdot \cos Dec}$$

Using this formula, we can find the approximate time for our observations (in analogy to sunrise and sunset).

As mentioned above, the simultaneous observation of stars **and** the horizon is possible during a limited time interval only.

To calculate the length of this interval, ΔT, we use the altitude formula and differentiate sin H with respect to the meridian angle, t:

$$\frac{d(\sin H)}{dt} = -\cos Lat \cdot \cos Dec \cdot \sin t$$

$$d(\sin H) = -\cos Lat \cdot \cos Dec \cdot \sin t \cdot dt$$

Substituting cosH•dH for d(sinH) and solving for dt, we get the change in the meridian angle, dt, as a function of a change in altitude, dH:

$$dt = -\frac{\cos H}{\cos Lat \cdot \cos Dec \cdot \sin t} \cdot dH$$

With H = –6° and dH = 6° (H = –3°...–9°), we get:

$$\Delta t[°] \approx -\frac{5.97}{\cos Lat \cdot \cos Dec \cdot \sin t}$$

Converting the change in the meridian angle to a time span (measured in minutes) and ignoring the sign, the equation is stated as:

$$\Delta T[m] \approx \frac{24}{\cos Lat \cdot \cos Dec \cdot \sin t}$$

The shortest possible time interval for our observations (Lat = 0, Dec = 0, t = 96°) lasts approx. 24 minutes. As the observer moves northward or southward from the equator, cos Lat and sin t decrease (t>90°). Accordingly, the duration of twilight increases. When t is 0° or 180°, ΔT is infinite.

This is in accordance with the well-known fact that twilight is shortest in equatorial regions and longest in polar regions.

We would obtain the same result when calculating t for H = –3° and H = –9°, respectively:

$$\Delta T[m] = 4 \cdot \left(t_{-9°}[°] - t_{-3°}[°]\right)$$

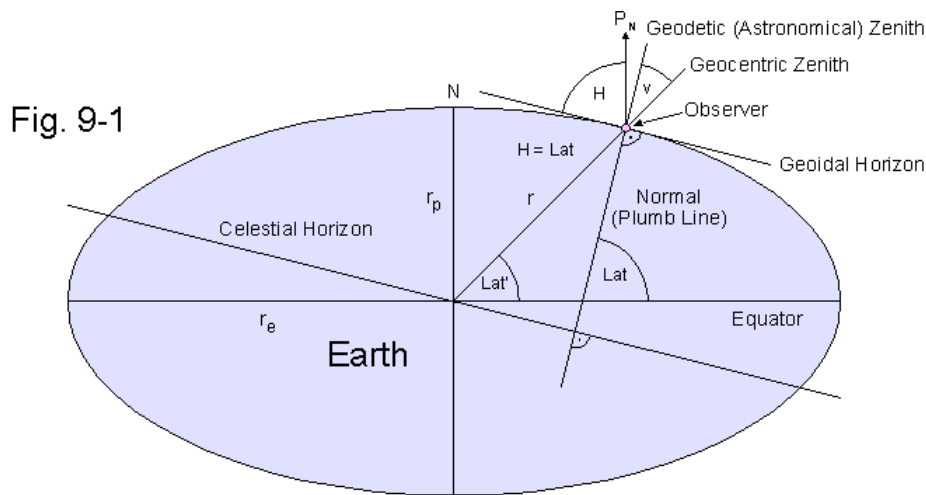
The Nautical Almanac provides tabulated values for the times of sunrise, sunset, civil twilight and nautical twilight for latitudes between –60° and +72° (referring to an observer being at the Greenwich meridian). In addition, times of moonrise and moonset are given.

Chapter 9

Geodetic Aspects of Celestial Navigation

The Ellipsoid

Celestial navigation is based upon the assumption that the earth is a sphere and, consequently, on the laws of spherical trigonometry. In reality, the shape of the earth is rather irregular and approximates an **oblate spheroid (ellipsoid)** resulting from two forces, **gravitation** and **centrifugal force**, acting on the viscous body of the earth. While gravitation alone would force the earth to assume the shape of a sphere, the state of lowest potential energy, the centrifugal force caused by the earth's rotation contracts the earth along the axis of rotation (polar axis) and stretches it along the plane of the equator. The local vector sum of both forces is called **gravity**.

A number of reference ellipsoids are in use to describe the shape of the earth, for example the **World Geodetic System (WGS)** ellipsoid of 1984. **The following considerations refer to the ellipsoid model of the earth which is sufficient for most navigational purposes**. *Fig.9-1* shows a meridional section of the ellipsoid.

Earth data (WGS 84 ellipsoid) :

Equatorial radius	r_e	6378137.0 m
Polar radius	r_p	6356752.3142 m
Flattening	$f = (r_e - r_p) / r_e$	1/298.25722

Due to the flattening of the earth, we have to distinguish between **geodetic** and **geocentric** latitude which would be the same if the earth were a sphere. The geodetic latitude of a given position, Lat, is the angle formed by the local normal (perpendicular) to the surface of the reference ellipsoid and the plane of the equator. The geocentric latitude, Lat', is the angle formed by the local radius vector and the plane of the equator. Geodetic and geocentric latitude are interrelated as follows:

$$\tan Lat' = (1-f)^2 \cdot \tan Lat$$

Geodetic and geocentric latitude are equal at the poles and on the equator. At all other places, the geocentric latitude, Lat', is smaller than the geodetic latitude, Lat. As with the spherical earth model, geodetic and geocentric longitude are equal. Maps are always based upon **geodetic coordinates.** These are also referred to as **geographic coordinates**.

In the following, we will discuss the effects of the oblateness (flattening) of the earth on celestial navigation.

Zenith distances (and altitudes) measured by the navigator always refer to the local plumb line which aligns itself with gravity and points to the **astronomical zenith**. Even the visible sea horizon correlates with the astronomical zenith since the water surface is perpendicular to the local plumb line.

With a homogeneous mass distribution throughout the ellipsoid, the plumb line coincides with the local normal to the ellipsoid which points to the **geodetic zenith**. Thus, astronomical and geodetic zenith are identical in this case.

The **geocentric zenith** is defined as the point where the extended local radius vector of the earth intersects the celestial sphere. The angular distance of the geodetic zenith from the geocentric zenith is called **angle of the vertical**, **v**. The angle of the vertical is a function of the geodetic latitude. The following formula was proposed by *Smart* [9]:

$$v[''] \approx 692.666 \cdot \sin(2 \cdot Lat) - 1.163 \cdot \sin(4 \cdot Lat) + 0.026 \cdot \sin(6 \cdot Lat)$$

The coefficients of the above formula refer to the proportions of the WGS 84 ellipsoid.

The angle of the vertical at a given position equals the difference between geodetic and geocentric latitude (*Fig. 9-1*):

$$v = Lat - Lat'$$

The maximum value of v, occurring at 45° geographic latitude, is approx. 11.5'. Thus, the geocentric latitude of an observer being at 45°geodetic latitude is only 44° 48.5'. This difference is not negligible. Therefore, the navigator has to know if the coordinates of a fix obtained by astronomical observations are geodetic or geocentric. Altitudes are measured with respect to the sea horizon or an artificial horizon. Both correlate with the local plumb line which points to the geodetic (astronomical) zenith. Thus, the latter is the only reference available to the navigator. As demonstrated in *Fig. 9-1*, the altitude of the celestial north pole, P_N, (corrected altitude of Polaris) with respect to the geoidal horizon equals the geodetic, not the geocentric latitude. A noon latitude, being the sum or difference of the (geocentric) declination and the zenith distance with respect to the geodetic zenith would give the same result.

Assuming a homogeneous mass distribution within the (ellipsoidal) earth, latitudes obtained by celestial observations are geodetic latitudes since the navigator measures altitudes with respect to the local geodetic zenith (directly or indirectly).

It is further important to know if the oblateness of the earth causes significant errors due to the fact that calculations of celestial navigation are based on a spherical earth model. According to the above values for polar radius and equatorial radius of the earth, the great circle distance of 1' is 1.849 km at the poles and 1.855 km at the equator. This small difference does not produce a significant error when plotting lines of position. It is therefore sufficient to use the adopted mean value (1 nautical mile = 1.852 km). However, when calculating the great circle distance (see chapter 11) of two locations thousands of nautical miles apart, the error caused by the oblateness of the earth can increase to several nautical miles. If extraordinary precision is required, the formulas for geodetic distance given in [2] should be used. A g**eodesic line** is the shortest path between two points on the surface of an ellipsoid. On the surface of a sphere, a geodesic line is the arc of a great circle.

The Parallax of the Moon

Due to the oblateness of the earth, the distance between geoidal and celestial horizon is not constant but can assume any value between r_p and r_e, depending on the observer's latitude. This has a measurable effect on the parallax of the moon since tabulated values for HP refer to the equatorial radius, r_e. The parallax of the moon is further affected by the displacement of the plumb line from the earth's center. A correction formula compensating both effects is given in chapter 2. The asymmetry of the plumb line with respect to the earth's center even causes a small (negligible) parallax in azimuth unless the moon is on the local meridian. In the following, we will calculate the effects of the oblateness of the earth on the parallax of the moon with the exact formulas of spherical astronomy [9]. For practical navigation, the simplified correction formulas given in chapter 2 are accurate enough.

Fig. 9-2 shows a projection of the astronomical zenith, Z_a, the geocentric zenith, Z_c, and the geographic position of the moon, M, on the **celestial sphere**, an imaginary hollow sphere of infinite diameter with the earth at its center.

The geocentric zenith, Z_c, is the point where a straight line from the earth's center through the observer's position intersects the celestial sphere. The astronomical zenith, Z_a, is the point at which the plumb line going through the observer's position intersects the celestial sphere. Z_a and Z_c are on the same meridian. M is the projected geocentric position of the moon defined by Greenwich hour angle and declination. Unfortunately, the position of a body defined by GHA and Dec is commonly called geographic position (see chapter 3) although GHA and Dec are geocentric coordinates. M' is the point where a straight line from the observer through the moon's center intersects the celestial sphere. Z_c, M, and M' are on a great circle. The zenith distance measured by the observer is z_a' because the astronomical zenith is the available reference. The quantity we want to know is z_a, the astronomical zenith distance corrected for parallax in altitude. This is the angular distance of the moon from the astronomical zenith, measured by a fictitious observer at the earth's center.
The known quantities are v, A_a', and z_a'. In contrast to the astronomer, the navigator is usually not able to measure A_a' precisely. For navigational purposes, the calculated azimuth (see chapter 4) may be substituted for A_a'.

We have three spherical triangles, Z_aZ_cM', Z_aZ_cM, and Z_aMM'. First, we calculate z_c' from z_a', v, and A_a' using the **law of cosines for sides** (see chapter 10):

$$\cos z_c' = \cos z_a' \cdot \cos v + \sin z_a' \cdot \sin v \cdot \cos\left(180° - A_a'\right)$$

$$z_c' = \arccos\left(\cos z_a' \cdot \cos v - \sin z_a' \cdot \sin v \cdot \cos A_a'\right)$$

To obtain z_c, we first have to calculate the relative length ($r_e = 1$) of the radius vector, r, and the geocentric parallax, p_c:

$$p_c = \arcsin\left(\rho \cdot \sin HP \cdot \sin z_c'\right)$$

$$\rho = \frac{r}{r_e} = \sqrt{\frac{1-\left(2e^2 - e^4\right) \cdot \sin^2 Lat}{1 - e^2 \cdot \sin^2 Lat}} \qquad e^2 = 1 - \frac{r_p^2}{r_e^2}$$

HP is the equatorial horizontal parallax. The geocentric zenith distance corrected for parallax is:

$$z_c = z_c' - p_c$$

Using the cosine formula again, we calculate A_c, the azimuth angle of the moon with respect to the geocentric zenith:

$$A_c = \arccos \frac{\cos z_a' - \cos z_c' \cdot \cos v}{\sin z_c' \cdot \sin v}$$

The astronomical zenith distance corrected for parallax is:

$$z_a = \arccos \left(\cos z_c \cdot \cos v + \sin z_c \cdot \sin v \cdot \cos A_c \right)$$

Thus, the **parallax in altitude** (astronomical) is:

$$PA = z_a' - z_a$$

For celestial navigation, the exact formulas of spherical astronomy are not needed, and the correction formula given in chapter 2 is accurate enough.

The small angle between M and M', measured at Z_a, is the **parallax in azimuth**, p_{az}:

$$p_{az} = \arccos \frac{\cos p_c - \cos z_a \cdot \cos z_a'}{\sin z_a \cdot \sin z_a'}$$

The correction for p_{az} is always applied so as to increase the angle formed by the azimuth line and the local meridian. For example, if Az_N is 315°, p_{az} is subtracted, and p_{az} is added if Az_N is 225°.

There is a simple formula for calculating the approximate parallax in azimuth:

$$p_{az} \approx f \cdot HP \cdot \frac{\sin(2 \cdot Lat) \cdot \sin Az_N}{\cos Hc}$$

This formula always returns the correct sign for p_{az}, and p_{az} is simply added to Az_N.

The parallax in azimuth does not exist when the moon is on the local meridian and is greatest when the moon ist east or west of the observer. It is further greatest at medium latitudes (45°) and non-existant when the observer is at one of the poles or on the equator (v = 0). Usually, the parallax in azimuth is only a fraction of an arcminute and therefore insignificant to celestial navigation. The parallax in azimuth increases with decreasing zenith distance.

Other celestial bodies do not require a correction for the oblateness of the earth since their parallaxes are very small compared with the parallax of the moon.

The Geoid

The earth is not **exactly** an oblate ellipsoid. The shape of the earth is more accurately described by the **geoid**, an equipotential surface of gravity.

The geoid has elevations and depressions caused by geographic features and a non-uniform mass distribution (materials of different density).

Elevations occur at local accumulations of matter (mountains, ore deposits), depressions at local deficiencies of matter (valleys, lakes, caverns). The elevation or depression of each point of the geoid with respect to the reference ellipsoid is found by gravity measurement.

On the slope of an elevation or depression of the geoid, the plumb line (the normal to the geoid) does not coincide with the normal to the reference ellipsoid, and the astronomical zenith differs from the geodetic zenith. Thus, an **astronomical position** (obtained through uncorrected astronomical observations) may slightly differ from the geodetic position. The small angle formed by the local plumb line and the local normal to the reference ellipsoid is called **deflection of the vertical**. Usually, this angle is smaller than one arcminute, but greater deflections of the vertical have been reported, for example, in coastal waters with adjacent high mountains.

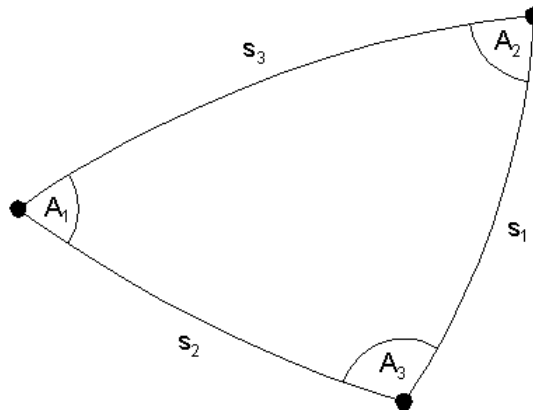
The local deflection of the vertical can be expressed in a latitude component and a longitude component. A position found by astronomical observations has to be corrected for both quantities to obtain the geodetic (geographic) position. The position error caused by the local deflection of the vertical is usually not relevant to celestial navigation at sea but is important to surveying and map-making where a much higher degree of accuracy is required.

Chapter 10

Spherical Trigonometry

The earth is usually regarded as a sphere in celestial navigation although an oblate spheroid would be a better approximation. Otherwise, navigational calculations would become too difficult for practical use. The position error introduced by the spherical earth model is usually very small and stays within the "statistical noise" caused by other omnipresent errors like, e.g., abnormal refraction, rounding errors, etc. Although it is possible to perform navigational calculations solely with the aid of tables (H.O. 229, H.O. 211, etc.) and with little mathematics, the principles of celestial navigation can not be comprehended without knowing the elements of spherical trigonometry.

The Oblique Spherical Triangle

Like any triangle, a spherical triangle is characterized by three sides and three angles. However, a spherical triangle is part of the surface of a sphere, and the sides are not straight lines but arcs of great circles (*Fig. 10-1*).

A **great circle** is a circle on the surface of a sphere whose plane passes through the center of the sphere (see chapter 3).

Any side of a spherical triangle can be regarded as an angle - the angular distance between the adjacent vertices, measured at the center of the sphere. The interrelations between angles and sides of a spherical triangle are described by the **law of sines**, the **law of cosines for sides**, the **law of cosines for angles**, **Napier's analogies**, and **Gauss' formulas** (apart from other formulas).

Law of sines:

$$\frac{\sin A_1}{\sin s_1} = \frac{\sin A_2}{\sin s_2} = \frac{\sin A_3}{\sin s_3}$$

Law of cosines for sides:

$$\cos s_1 = \cos s_2 \cdot \cos s_3 + \sin s_2 \cdot \sin s_3 \cdot \cos A_1$$
$$\cos s_2 = \cos s_1 \cdot \cos s_3 + \sin s_1 \cdot \sin s_3 \cdot \cos A_2$$
$$\cos s_3 = \cos s_1 \cdot \cos s_2 + \sin s_1 \cdot \sin s_2 \cdot \cos A_3$$

Law of cosines for angles:

$$\cos A_1 = -\cos A_2 \cdot \cos A_3 + \sin A_2 \cdot \sin A_3 \cdot \cos s_1$$
$$\cos A_2 = -\cos A_1 \cdot \cos A_3 + \sin A_1 \cdot \sin A_3 \cdot \cos s_2$$
$$\cos A_3 = -\cos A_1 \cdot \cos A_2 + \sin A_1 \cdot \sin A_2 \cdot \cos s_3$$

Napier's analogies:

$$\tan\frac{A_1+A_2}{2}\cdot\tan\frac{A_3}{2} = \frac{\cos\frac{s_1-s_2}{2}}{\cos\frac{s_1+s_2}{2}} \qquad \tan\frac{A_1-A_2}{2}\cdot\tan\frac{A_3}{2} = \frac{\sin\frac{s_1-s_2}{2}}{\sin\frac{s_1+s_2}{2}}$$

$$\frac{\tan\frac{s_1+s_2}{2}}{\tan\frac{s_3}{2}} = \frac{\cos\frac{A_1-A_2}{2}}{\cos\frac{A_1+A_2}{2}} \qquad \frac{\tan\frac{s_1-s_2}{2}}{\tan\frac{s_3}{2}} = \frac{\sin\frac{A_1-A_2}{2}}{\sin\frac{A_1+A_2}{2}}$$

Gauss' formulas:

$$\frac{\sin\frac{A_1+A_2}{2}}{\cos\frac{A_3}{2}} = \frac{\cos\frac{s_1-s_2}{2}}{\cos\frac{s_3}{2}} \qquad \frac{\cos\frac{A_1+A_2}{2}}{\sin\frac{A_3}{2}} = \frac{\cos\frac{s_1+s_2}{2}}{\cos\frac{s_3}{2}}$$

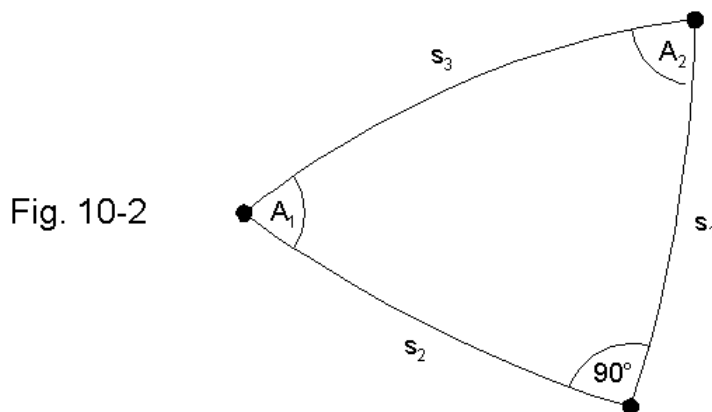
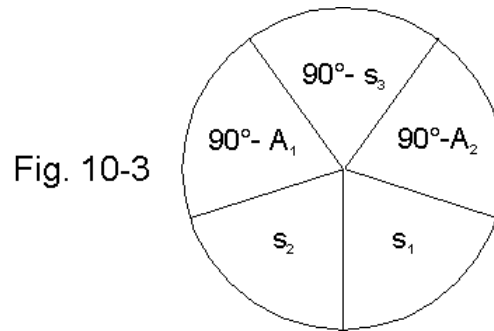
$$\frac{\sin\frac{A_1-A_2}{2}}{\cos\frac{A_3}{2}} = \frac{\sin\frac{s_1-s_2}{2}}{\sin\frac{s_3}{2}} \qquad \frac{\cos\frac{A_1-A_2}{2}}{\sin\frac{A_3}{2}} = \frac{\sin\frac{s_1+s_2}{2}}{\sin\frac{s_3}{2}}$$

These formulas and others derived thereof enable any quantity (angle or side) of a spherical triangle to be calculated if three other quantities are known.

Particularly the law of cosines for sides is of interest to the navigator.

The Right Spherical Triangle

Solving a spherical triangle is less complicated when it contains a right angle (*Fig. 10-2*). Using ***Napier's* rules of circular parts**, any quantity can be calculated if only two other quantities (apart from the right angle) are known.

Fig. 10-2

We arrange the sides forming the right angle (s_1, s_2) and the **complements** of the remaining angles (A_1, A_2) and opposite side (s_3) in the form of a circular diagram consisting of five sectors, called "parts" (in the same order as they appear in the triangle). The right angle itself is omitted (*Fig. 10-3*):

According to *Napier's* rules, the sine of any part of the diagram equals the product of the tangents of the adjacent parts and the product of the cosines of the opposite parts:

$$\sin s_1 = \tan s_2 \cdot \tan\left(90° - A_2\right) = \cos\left(90° - A_1\right) \cdot \cos\left(90° - s_3\right)$$
$$\sin s_2 = \tan\left(90° - A_1\right) \cdot \tan s_1 = \cos\left(90° - s_3\right) \cdot \cos\left(90° - A_2\right)$$
$$\sin\left(90° - A_1\right) = \tan\left(90° - s_3\right) \cdot \tan s_2 = \cos\left(90° - A_2\right) \cdot \cos s_1$$
$$\sin\left(90° - s_3\right) = \tan\left(90° - A_2\right) \cdot \tan\left(90° - A_1\right) = \cos s_1 \cdot \cos s_2$$
$$\sin\left(90° - A_2\right) = \tan s_1 \cdot \tan\left(90° - s_3\right) = \cos s_2 \cdot \cos\left(90° - A_1\right)$$

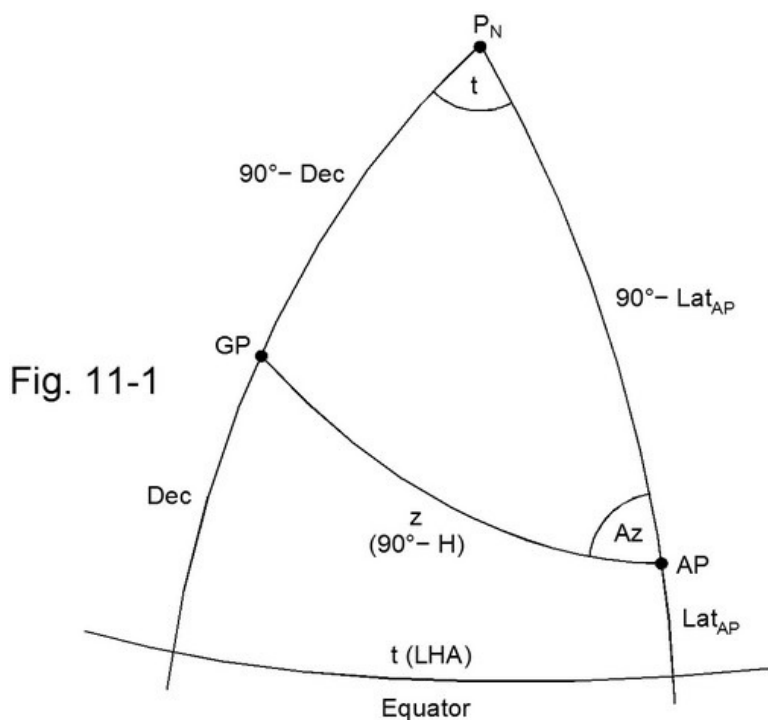
In a simpler form, these equations are stated as:

$$\sin s_1 = \tan s_2 \cdot \cot A_2 = \sin A_1 \cdot \sin s_3$$
$$\sin s_2 = \cot A_1 \cdot \tan s_1 = \sin s_3 \cdot \sin A_2$$
$$\cos A_1 = \cot s_3 \cdot \tan s_2 = \sin A_2 \cdot \cos s_1$$
$$\cos s_3 = \cot A_2 \cdot \cot A_1 = \cos s_1 \cdot \cos s_2$$
$$\cos A_2 = \tan s_1 \cdot \cot s_3 = \cos s_2 \cdot \sin A_1$$

There are several applications for the right spherical triangle in navigation, for example **Ageton's sight reduction tables** (chapter 11) and **great circle navigation** (chapter 13).

Chapter 11

The Navigational Triangle

The **navigational triangle** is the (usually) oblique spherical triangle formed by the north pole, P_N, the observer's assumed position, AP, and the geographic position of the celestial object, GP (*Fig. 11-1*). **All common sight reduction procedures are based upon the navigational triangle.**

When using the intercept method, the latitude of the assumed position, Lat_{AP}, the declination of the observed celestial body, Dec, and the meridian angle, t, or the local hour angle, LHA, (calculated from the longitude of AP and the GHA of the object), are the known quantities.

The first step is calculating the side z of the navigational triangle by using the law of cosines for sides:

$$\cos z = \cos\left(90° - Lat_{AP}\right) \cdot \cos\left(90° - Dec\right) + \sin\left(90° - Lat_{AP}\right) \cdot \sin\left(90° - Dec\right) \cdot \cos t$$

Since cos (90°-x) equals sin x and vice versa, the equation can be written in a simpler form:

$$\cos z = \sin Lat_{AP} \cdot \sin Dec + \cos Lat_{AP} \cdot \cos Dec \cdot \cos t$$

The side z is not only the **great circle distance** between AP and GP but also the **zenith distance** of the celestial object and the radius of the **circle of equal altitude** (see chapter 1).

Substituting the altitude H for z, we get:

$$\sin H = \sin Lat_{AP} \cdot \sin Dec + \cos Lat_{AP} \cdot \cos Dec \cdot \cos t$$

Solving the equation for H leads to the altitude formula known from chapter 4:

$$H = \arcsin\left(\sin Lat_{AP} \cdot \sin Dec + \cos Lat_{AP} \cdot \cos Dec \cdot \cos t\right)$$

The altitude thus obtained for a given position is called **computed altitude**, **Hc**.

The azimuth angle of the observed body is also calculated by means of the law of cosines for sides:

$$\cos(90° - Dec) = \cos(90° - Lat_{AP}) \cdot \cos z + \sin(90° - Lat_{AP}) \cdot \sin z \cdot \cos Az$$

$$\sin Dec = \sin Lat_{AP} \cdot \cos z + \cos Lat_{AP} \cdot \sin z \cdot \cos Az$$

Using the computed altitude instead of the zenith distance results in the following equation:

$$\sin Dec = \sin Lat_{AP} \cdot \sin Hc + \cos Lat_{AP} \cdot \cos Hc \cdot \cos Az$$

Solving the equation for Az finally yields the azimuth formula from chapter 4:

$$Az = \arccos \frac{\sin Dec - \sin Lat_{AP} \cdot \sin Hc}{\cos Lat_{AP} \cdot \cos Hc}$$

The arccos function returns angles between 0° and 180°. Therefore, the resulting azimuth angle is not necessarily identical with the true azimuth, Az_N (0°... 360°, measured clockwise from true north) commonly used in navigation. In all cases where t is negative (GP east of AP) , Az equals Az_N. Otherwise (t positive, GP westward from AP as shown in *Fig. 11-1*), Az_N is obtained by subtracting Az from 360°.

When the meridian angle, t, (or the local hour angle, LHA) is the quantity to be calculated (time sight, *Sumner's* method), Dec, Lat_{AP} (the assumed latitude), and z (or H) are the known quantities.

Again, the law of cosines for sides is applied:

$$\cos z = \cos(90° - Lat_{AP}) \cdot \cos(90° - Dec) + \sin(90° - Lat_{AP}) \cdot \sin(90° - Dec) \cdot \cos t$$

$$\sin H = \sin Lat_{AP} \cdot \sin Dec + \cos Lat_{AP} \cdot \cos Dec \cdot \cos t$$

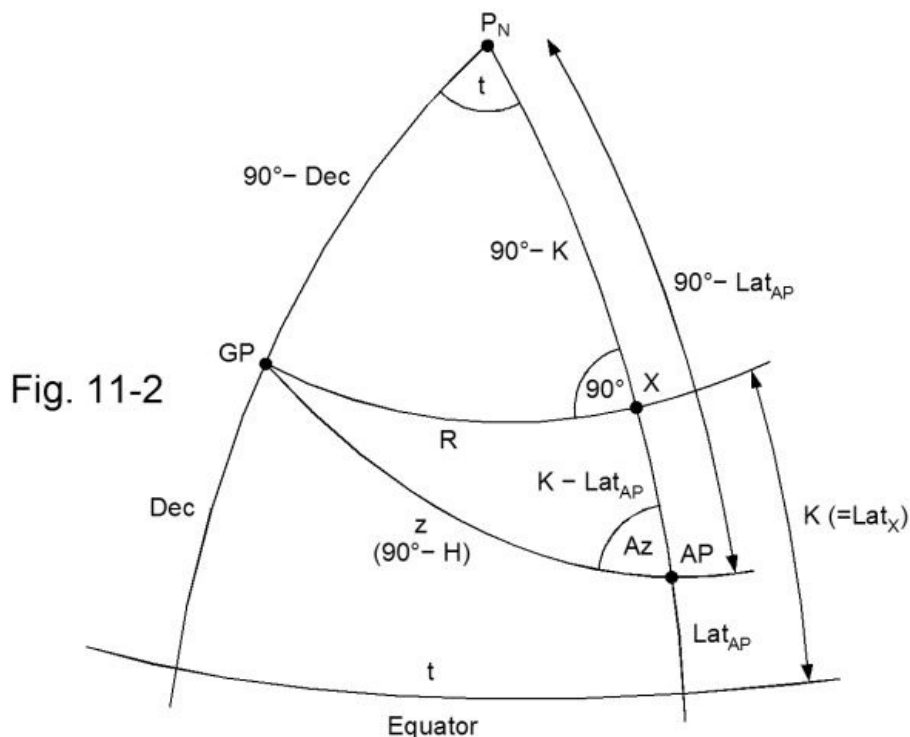
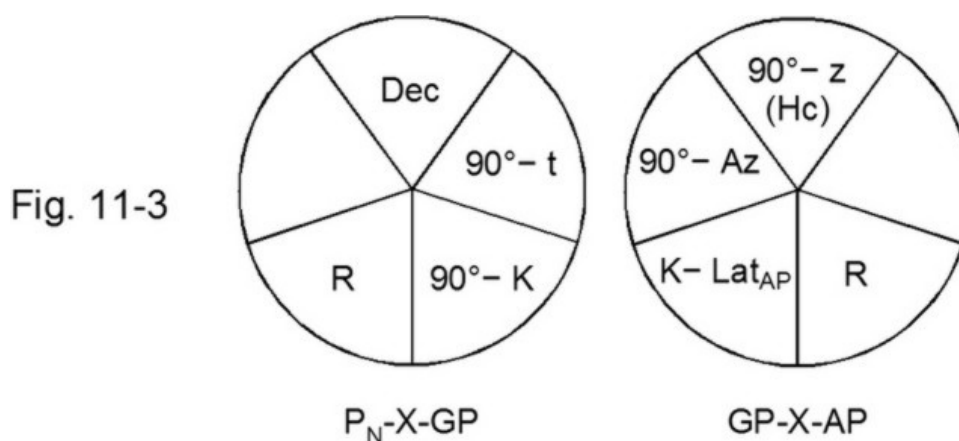
$$\cos t = \frac{\sin H - \sin Lat_{AP} \cdot \sin Dec}{\cos Lat_{AP} \cdot \cos Dec}$$

$$t = \arccos \frac{\sin H - \sin Lat_{AP} \cdot \sin Dec}{\cos Lat_{AP} \cdot \cos Dec}$$

The obtained meridian angle, t (or LHA), is then used as described in chapter 4 and chapter 5.

When observing a celestial body at the time of meridian passage (e. g., for determining one's latitude), the local hour angle is zero, and the navigational triangle becomes infinitesimally narrow. In this case, the formulas of spherical trigonometry are not needed, and the sides of the spherical triangle can be calculated by simple addition or subtraction.

The Divided Navigational Triangle

An alternative method for solving the navigational triangle is based upon two right spherical triangles obtained by constructing a great circle passing through GP and intersecting the local meridian perpendicularly at X (*Fig. 11-2*):

The first right triangle is formed by P_N, X, and GP, the second one by GP, X, and AP. The **auxiliary parts** R and K are intermediate quantities used to calculate z (or Hc) and Az. K is the geographic latitude of X. Both triangles are solved using **Napier's rules of circular parts** (see chapter 9). *Fig. 11-3* illustrates the corresponding circular diagrams:

According to Napier's rules, Hc and Az are calculated by means of the following formulas:

$$\sin R = \sin t \cdot \cos Dec \qquad \Rightarrow \qquad R = \arcsin\left(\sin t \cdot \cos Dec\right)$$

$$\sin Dec = \cos R \cdot \sin K \qquad \Rightarrow \qquad \sin K = \frac{\sin Dec}{\cos R} \qquad \Rightarrow \qquad K = \arcsin \frac{\sin Dec}{\cos R}$$

Substitute 180°–K for K in the following equation if $|t| > 90°$ (or $90° < LHA < 270°$).

$$\sin Hc = \cos R \cdot \cos\left(K - Lat_{AP}\right) \quad \Rightarrow \quad Hc = \arcsin\left[\cos R \cdot \cos\left(K - Lat_{AP}\right)\right]$$

$$\sin R = \cos Hc \cdot \sin Az \quad \Rightarrow \quad \sin Az = \frac{\sin R}{\cos Hc} \quad \Rightarrow \quad Az = \arcsin \frac{\sin R}{\cos Hc}$$

For further calculations, substitute 180°−Az for Az if K and Lat have opposite signs or if $|K| < |Lat|$.

To obtain the true azimuth, Az_N (0°... 360°), the following rules have to be applied:

$$Az_N = \begin{cases} -Az & \text{if} \quad Lat_{AP} > 0\,(\text{N}) \quad \text{AND} \quad t < 0 \ \left(180° < LHA < 360°\right) \\ 360° - Az & \text{if} \quad Lat_{AP} > 0\,(\text{N}) \quad \text{AND} \quad t > 0 \ \left(0° < LHA < 180°\right) \\ 180° + Az & \text{if} \quad Lat_{AP} < 0\,(\text{S}) \end{cases}$$

The divided navigational triangle is of considerable importance since it forms the theoretical background for a number of **sight reduction tables**, e.g., the Ageton Tables (see below). It is also used for **great circle navigation** (chapter 12).

Using the secant and cosecant function (sec x = 1/cos x, csc x = 1/sin x), we can write the equations for the divided navigational triangle in the following form:

$$\csc R = \csc t \cdot \sec Dec$$

$$\csc K = \frac{\csc Dec}{\sec R}$$

Substitute 180°−K for K in the following equation if $|t| > 90°$:

$$\csc Hc = \sec R \cdot \sec\left(K - Lat\right)$$

$$\csc Az = \frac{\csc R}{\sec Hc}$$

Substitute 180°−Az for Az if K and Lat have opposite signs or if $|K| < |Lat|$.

In logarithmic form, these equations are stated as:

$$\log\csc R = \log\csc t + \log\sec Dec$$

$$\log\csc K = \log\csc Dec - \log\sec R$$

$$\log\csc Hc = \log\sec R + \log\sec\left(K - Lat\right)$$

$$\log\csc Az = \log\csc R - \log\sec Hc$$

With the logarithms of the secants and cosecants of angles arranged in the form of a suitable table, we can solve a sight by a sequence of simple additions and subtractions. Apart from the table itself, the only tools required are a sheet of paper and a pencil.

The Ageton Tables (H.O. 211), first published in 1931, are based upon the above formulas and provide a very efficient arrangement of angles and their log secants and log cosecants on 36 pages. Since all calculations are based on absolute values, certain rules included in the instructions have to be observed.

Sight reduction tables were developed many years before electronic calculators became available in order to simplify calculations necessary to reduce a sight. Still today, sight reduction tables are preferred by people who do not want to deal with the formulas of spherical trigonometry. Moreover, they provide a valuable backup method if electronic devices fail.

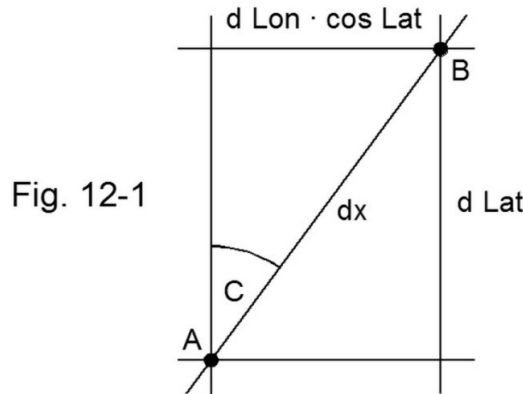
Two modified versions of the Ageton Tables are available at: http://www.celnav.de/page3.htm

Chapter 12

General Formulas for Navigation

Although the following formulas are not part of celestial navigation, they are indispensible because they are necessary to calculate distance and direction (course) from the **point of departure**, A, to the **point of arrival**, B, as well as to calculate the position of B from the position of A if course and distance are known. The **true course**, C, is the angle made by the vector of motion and the local meridian. It is measured from true north (clockwise through 360°). Knowing the coordinates of A, Lat_A and Lon_A, and the coordinates of B, Lat_B and Lon_B, the navigator has the principal choice between **rhumb line navigation** (simple procedure but longer distance) and **great circle navigation** (shortest possible distance on a sphere). Combinations of both methods are possible.

Rhumb Line Navigation

A **rhumb line** (also called **loxodrome**) is a line on the surface of the earth intersecting all meridians at a constant angle, C. Thus, a rhumb line is represented by a straight line on a Mercartor chart (see chapter 13) which makes **voyage planning** quite simple. On a globe, a rhumb line forms a spherical spiral extending from pole to pole unless it is identical with a meridian (C = 0° or 180°) or a parallel of latitude (C = 90° or 270°). A vessel steering a constant course travels along a rhumb line, provided there is no drift. Rhumb line course, C, and distance, d, are calculated as shown below. First, we imagine traveling the infinitesimal distance dx from the point of departure, A, to the point of arrival, B. Our course is C (*Fig. 12-1*):

The distance, dx, is the vector sum of a north-south component, dLat, and a west-east component, dLon · cos Lat. The factor cos Lat is the relative circumference of the respective parallel of latitude (equator = 1):

$$\tan C = \frac{d\,Lon \cdot \cos Lat}{d\,Lat}$$

$$\frac{d\,Lat}{\cos Lat} = \frac{1}{\tan C} \cdot d\,Lon$$

If the distance between A (defined by Lat_A and Lon_A) and B (defined by Lat_B and Lon_B) is a measurable quantity, we have to integrate:

$$\int_{Lat\,A}^{Lat\,B} \frac{d\,Lat}{\cos Lat} = \frac{1}{\tan C} \cdot \int_{Lon\,A}^{Lon\,B} d\,Lon$$

$$\ln\left[\tan\left(\frac{Lat_B}{2}+\frac{\pi}{4}\right)\right]-\ln\left[\tan\left(\frac{Lat_A}{2}+\frac{\pi}{4}\right)\right]=\frac{Lon_B-Lon_A}{\tan C}$$

$$\tan C=\frac{Lon_B-Lon_A}{\ln\dfrac{\tan\left(\dfrac{Lat_B}{2}+\dfrac{\pi}{4}\right)}{\tan\left(\dfrac{Lat_A}{2}+\dfrac{\pi}{4}\right)}}$$

Solving for C and measuring angles in degrees, we get:

$$C=\arctan\frac{Lon_B-Lon_A}{\ln\dfrac{\tan\left(\dfrac{Lat_B}{2}+45°\right)}{\tan\left(\dfrac{Lat_A}{2}+45°\right)}}$$

The term Lon_B-Lon_A has to be in the range between –180° tand +180°. If it is outside this range, we have to add or subtract 360° before entering the rhumb line course formula.

The arctan function returns values between -90° and +90°. To obtain the true course (0°...360°), we apply the following rules:

$$C\rightarrow\begin{cases}C & \text{if}\quad Lat_B>Lat_A\quad\text{AND}\quad Lon_B>Lon_A\\360°+C & \text{if}\quad Lat_B>Lat_A\quad\text{AND}\quad Lon_B<Lon_A\\180°+C & \text{if}\quad Lat_B<Lat_A\end{cases}$$

To find the total length of the rhumb line track, we calculate the infinitesimal distance dx:

$$d\,x=\frac{d\,Lat}{\cos C}$$

The total length is found through integration:

$$d=\frac{1}{\cos C}\cdot\int_{Lat\,A}^{Lat\,B}d\,Lat=\frac{Lat_B-Lat_A}{\cos C}$$

Finally, we get:

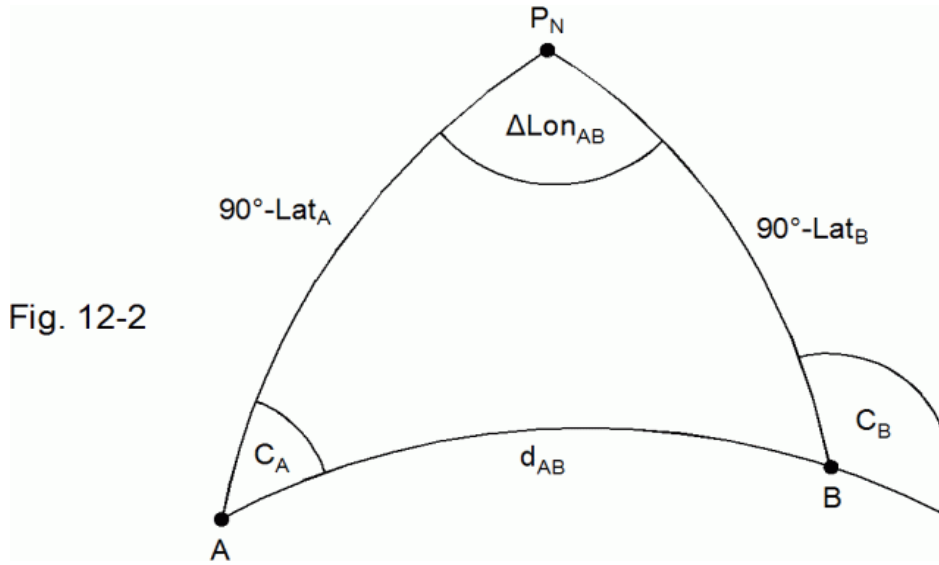
$$d\,[km]=\frac{40031.6}{360}\cdot\frac{Lat_B-Lat_A}{\cos C}\qquad d\,[nm]=60\cdot\frac{Lat_B-Lat_A}{\cos C}$$

If both positions have the same latitude, the distance can **not** be calculated using the above formulas. In this case, the following formulas apply (C is either 90° or 270°.):

$$d\,[km] = \frac{40031.6}{360} \cdot \left(Lon_B - Lon_A \right) \cdot \cos Lat \qquad d\,[nm] = 60 \cdot \left(Lon_B - Lon_A \right) \cdot \cos Lat$$

Great Circle Navigation

Great circle distance, d_{AB}, and course, C_A, are calculated on the analogy of zenith distance and azimuth. For this pupose, we consider the navigational triangle (see chapter 11) and substitute A for GP, B for AP, d_{AB} for z, and ΔLon_{AB} (difference of longitude) for LHA (*Fig. 12-2*):

Fig. 12-2

$$d_{AB} = \arccos\left(\sin Lat_A \cdot \sin Lat_B + \cos Lat_A \cdot \cos Lat_B \cdot \cos \Delta Lon_{AB} \right)$$

$$\Delta Lon_{AB} = Lon_B - Lon_A$$

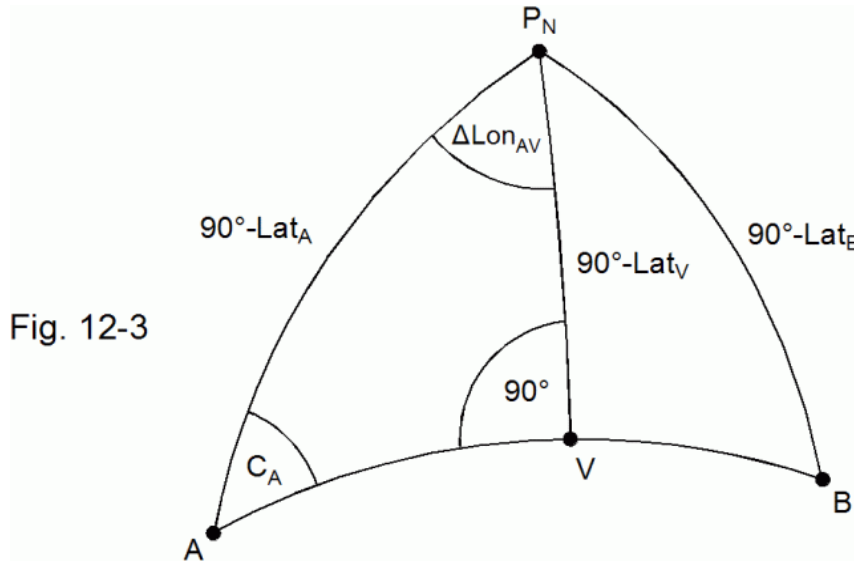
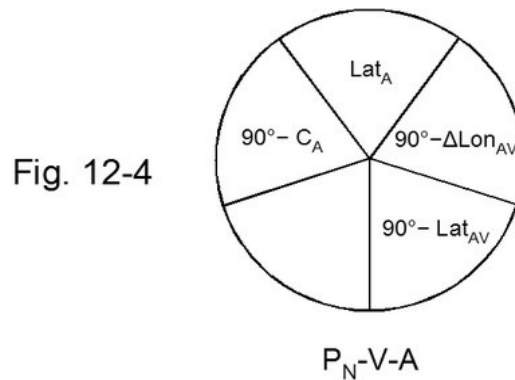
Northern latitude and eastern longitude are positive, southern latitude and western longitude negative. A great circle distance has the dimension of an angle. To measure it in distance units, we multiply it by 40031.6/360 (distance in km) or by 60 (distance in nm).

$$C_A = \arccos \frac{\sin Lat_B - \sin Lat_A \cdot \cos d_{AB}}{\cos Lat_A \cdot \sin d_{AB}}$$

If the term sin (Lon_B-Lon_A) is negative, we replace C_A with 360°-C_A in order to obtain the true course (0°... 360° clockwise from true north).

In *Fig. 12-2*, C_A is the **initial great circle course**, C_B the **final great circle course**. Since the angle between the great circle and the respective local meridian varies as we progress along the great circle (unless the great circle coincides with the equator or a meridian), we can **not** steer a constant course as we would when following a rhumb line.
Theoretically, we have to adjust the course continually. This is possible with the aid of navigation computers and autopilots. If such means are not available, we have to calculate an updated course at certain intervals (see below).

Great circle navigation requires more careful voyage planning than rhumb line navigation. On a Mercator chart (see chapter 13), a great circle track appears as a line bent towards the equator. As a result, the navigator may need more information about the intended great circle track in order to verify if it leads through navigable areas.

With the exception of the equator, every great circle has two **vertices**, the points farthest from the equator. The vertices have the same absolute value of latitude (with opposite sign) but are 180° apart in longitude. At each **vertex** (also called **apex**), the great circle is tangent to a parallel of latitude, and C is either 90° or 270° (cos C = 0). Thus, we have a right spherical triangle formed by the north pole, P_N, the vertex,V, and the point of departure, A (Fig. 12-3):

To derive the formulas needed for the following calculations, we use Napier's rules of circular parts (Fig. 12-4). The right angle is at the bottom of the circular diagram. The five parts are arranged clockwise.

First, we need the latitude of the vertex, Lat_V:

$$\cos Lat_V = \sin C_A \cdot \cos Lat_A$$

Solving for Lat_V, we get:

$$Lat_V = \pm \arccos\left(\left|\sin C_A\right| \cdot \cos Lat_A\right)$$

The absolute value of sin C_A is used to make sure that Lat_V does not exceed ±90° (the arccos function returns values between 90° and 180° for negative arguments). The equation has two solutions, according to the number of vertices. Only the vertex lying ahead of us is relevant to voyage planning. It is found using the following modified formula:

$$Lat_V = \operatorname{sgn}\left(\cos C_A\right) \cdot \arccos\left(\left|\sin C_A\right| \cdot \cos Lat_A\right)$$

sng(x) is the **signum** function:

$$\operatorname{sng}(x) = \begin{cases} -1 & \text{if} \quad x < 0 \\ 0 & \text{if} \quad x = 0 \\ +1 & \text{if} \quad x > 0 \end{cases}$$

If V is located between A and B (like shown in Fig. 12-3), our latitude passes through an extremum at the instant we reach V. This does not happen if B is between A and V.

Knowing Lat_V, we are able to calculate the longitude of V. Again, we apply Napier's rules:

$$\cos \Delta Lon_{AV} = \frac{\tan Lat_A}{\tan Lat_V} \qquad \Delta Lon_{AV} = \Delta Lon_V - \Delta Lon_A$$

Solving for ΔLon_{AV}, we get:

$$\Delta Lon_{AV} = \arccos \frac{\tan Lat_A}{\tan Lat_V}$$

The longitude of V is

$$Lon_V = Lon_A + \operatorname{sgn}\left(\sin C_A\right) \cdot \arccos \frac{\tan Lat_A}{\tan Lat_V}$$

(Add or subtract 360° if necessary.)

The term sng (sin C_A) in the above formula provides an automatic correction for the sign of ΔLon_{AV}.

Knowing the position of V (defined by Lat_V and Lon_V), we are now able to calculate the position of any chosen point, X, on the intended great circle track (substituting X for A in the right spherical triangle). Using Napier's rules once more, we get:

$$\tan Lat_X = \cos \Delta Lon_{XV} \cdot \tan Lat_V \qquad \Delta Lon_{XV} = Lon_V - Lon_X$$

$$Lat_X = \arctan \left(\cos \Delta Lon_{XV} \cdot \tan Lat_V\right)$$

Further, we can calculate the course at the point X:

$$\cos C_X = \sin \Delta Lon_{XV} \cdot \sin Lat_V$$

$$C_X = \begin{cases} \arccos\left(\sin \Delta Lon_{XV} \cdot \sin Lat_V\right) & \text{if} \quad \sin C_A > 0 \\ 180° + \arccos\left(\sin \Delta Lon_{XV} \cdot \sin Lat_V\right) & \text{if} \quad \sin C_A < 0 \end{cases}$$

Alternatively, C_X can be calculated from the oblique spherical triangle formed by X, P_N, and B.

The above formulas enable us to establish suitably spaced **waypoints** on the great circle and connect them by straight lines on the Mercator chart. The series of **legs** thus obtained, each one being a rhumb line track, is a practical approximation of the intended great circle track. Further, we are now able to see beforehand if there are obstacles in our way.

Mean latitude

Because of their simplicity, the **mean latitude** formulas are often used in everyday navigation. Mean latitude is a good approximation for rhumb line navigation for short and medium distances between A and B. The method is less suitable for polar regions (convergence of meridians).

Course:

$$C = \arctan\left(\cos Lat_M \cdot \frac{Lon_B - Lon_A}{Lat_B - Lat_A}\right) \qquad Lat_M = \frac{Lat_A + Lat_B}{2}$$

The true course is obtained by applying the same rules to C as to the rhumb line course (see above).
Distance:

$$d\,[km] = \frac{40031.6}{360} \cdot \frac{Lat_B - Lat_A}{\cos C} \qquad d\,[nm] = 60 \cdot \frac{Lat_B - Lat_A}{\cos C}$$

If C = 90° or C = 270°, we have use the following formulas:

$$d\,[km] = \frac{40031.6}{360} \cdot \left(Lon_B - Lon_A\right) \cdot \cos Lat \qquad d\,[nm] = 60 \cdot \left(Lon_B - Lon_A\right) \cdot \cos Lat$$

Dead Reckoning

Dead reckoning is the navigational term for extrapolating one's new position, B, from the previous position, A, the course C, and the distance, d (calculated from the vessel's average speed and time elapsed). The position thus obtained is called a **dead reckoning position, DRP**.

Since a DRP is only an approximate position (due to the influence of drift, etc.), the mean latitude method (see above) provides sufficient accuracy. On land, dead reckoning is of limited use since it is usually not possible to steer a constant course (apart from driving in large, entirely flat areas like, e.g., salt flats).

At sea, the DRP is needed to choose a suitable AP for the intercept method. If celestial observations are not possible and electronic navigation aids not available, dead reckoning may be the only way of keeping track of one's position. Apart from the very simple graphic solution, there are two formulas for the calculation of the DRP.

Calculation of new latitude:

$$Lat_B[°] = Lat_A[°] + \frac{360}{40031.6} \cdot d\,[km] \cdot \cos C \quad \text{or} \quad Lat_B[°] = Lat_A[°] + \frac{d\,[nm] \cdot \cos C}{60}$$

Calculation of new longitude:

$$Lon_B[°] = Lon_A[°] + \frac{360}{40031.6} \cdot \frac{d\,[km] \cdot \sin C}{\cos Lat_M} \quad \text{or} \quad Lon_B[°] = Lon_A[°] + \frac{d\,[nm] \cdot \sin C}{60 \cdot \cos Lat_M}$$

If the resulting longitude is greater than +180°, we subtract 360°. If it is smaller than -180°, we add 360°.

If our movement is composed of several components (including drift, etc.), we have to replace the terms d·cos C and d·sin C with

$\sum d_i \cdot \cos C_i$ and $\sum d_i \cdot \sin C_i$, respectively.

Chapter 13

Charts and Plotting Sheets

Mercator Charts

Sophisticated navigation is not possible without the use of a map (chart), a projection of a certain area of the earth's surface with its geographic features on a plane. Among the numerous types of map projection, the **Mercator projection,** named after the Flemish-German cartographer Gerhard Kramer (Latin: Gerardus Mercator), is mostly used in navigation because it produces charts with an orthogonal grid which is most convenient for measuring directions and plotting lines of position. Further, rhumb lines appear as straight lines on a Mercator chart. Great circles do not, apart from meridians and the equator which are also rhumb lines.

In order to construct a Mercator chart, we have to remember how the grid printed on a globe looks. At the equator, an area of, e. g., 2 by 2 degrees looks almost like a square, but it appears as a narrow trapezoid when we place it near one of the poles. While the distance between two adjacent parallels of latitude is constant, the distance between two meridians becomes progressively smaller as the latitude increases because the meridians converge to the poles. An area with the infinitesimal dimensions dLat and dLon would appear as an oblong with the dimensions dx and dy on our globe (*Fig. 13-1*):

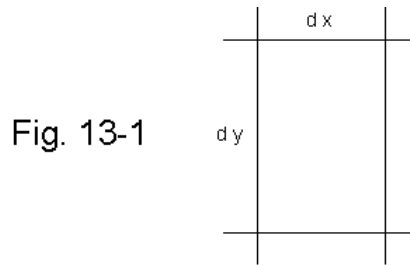



Fig. 13-1

$$dx = c' \cdot d\,Lon \cdot \cos Lat$$

$$dy = c' \cdot d\,Lat$$

dx contains the factor cos Lat since the circumference of a parallel of latitude is proportional to cos Lat. The constant c' is the scale factor of the globe (measured in, e. g., mm/°).

Since we require **any** rhumb line to appear as a **straight** line intersecting all meridians at a constant angle, meridians have to be equally spaced vertical lines on our chart, and any infinitesimal oblong defined by dLat and dLon must have the same aspect ratio as on the globe (dy/dx = const.) at a given latitude (**conformality**).

Therefore, if we transfer the oblong defined by dLat and dLon from the globe to our chart, we get the dimensions:

$$dx = c \cdot d\,Lon$$

$$dy = c \cdot \frac{d\,Lat}{\cos Lat}$$

The new constant c is the scale factor of the chart. Now, dx remains constant (parallel meridians), but dy is a function of the latitude at which our small oblong is located. To obtain the smallest distance from any point with the latitude Lat_P to the equator, we integrate:

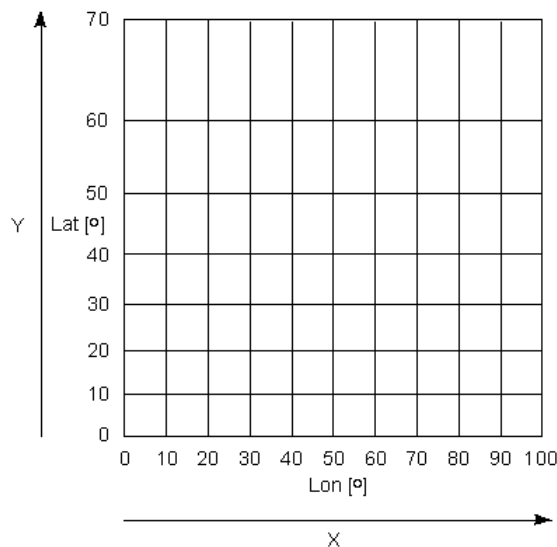
$$Y = \int_0^Y dy = c \cdot \int_0^{Lat_P} \frac{d\,Lat}{\cos Lat} = c \cdot \ln \tan\left(\frac{Lat_P}{2} + \frac{\pi}{4} \right)$$

Y is the distance of the respective parallel of latitude from the equator. In the above equation, angles are given in circular measure (radians). If we measure angles in degrees, the equation is stated as:

$$Y = c \cdot \ln \tan\left(\frac{Lat_P[°]}{2} + 45° \right)$$

The distance of any point from the Greenwich meridian (Lon = 0°) varies proportionally with the longitude of the point, Lon_P. X is the distance of the respective meridian from the Greenwich meridian:

$$X = \int_0^{Lon_P} dx = c \cdot Lon_P$$

Fig. 13-2 shows an example of the resulting **graticule** (10° spacing). While meridians of longitude appear as equally spaced vertical lines, parallels of latitude are horizontal lines drawn farther apart as the latitude increases. Y would be infinite at 90° latitude.

Mercator charts have the disadvantage that geometric distortions increase as the distance from the equator increases. The Mercator projection is therefore not suitable for polar regions. A circle of equal altitude, for example, would appear as a distorted ellipse at higher latitudes. Areas near the poles, e. g., Greenland, appear much greater on a Mercator map than on a globe.

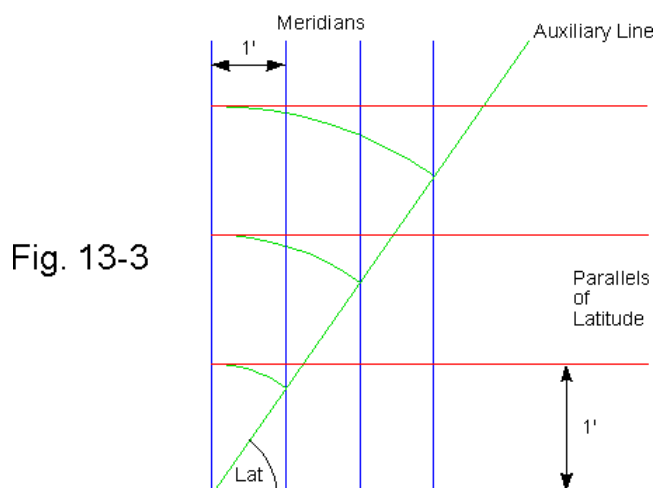
It is often said that a Mercator chart is obtained by projecting each point of the surface of a globe from the center of the globe to the inner surface of a hollow cylinder tangent to the globe at the equator. This is only a rough approximation. As a result of such a (purely geometrical) projection, Y would be proportional to tan Lat, and conformality would not be achieved.

Plotting Sheets

If we magnify a small part of a Mercator chart, e. g., an area of 30' latitude by 40' longitude, we will notice that the spacing between the parallels of latitude now seems to be almost constant. An approximated Mercator grid of such a small area can be constructed by drawing equally spaced horizontal lines, representing the parallels of latitude, and equally spaced vertical lines, representing the meridians.

The spacing of the parallels of latitude, Δy, defines the scale of our chart, e. g., 5mm/nm. The spacing of the meridians, Δx, is a function of the mean latitude, Lat_M:

$$\Delta x = \Delta y \cdot \cos Lat_M \qquad Lat_M = \frac{Lat_{min} + Lat_{max}}{2}$$

A sheet of otherwise blank paper with such a simplified Mercator grid is called a **small area plotting sheet** and is a very useful tool for plotting lines of position. If a calculator or trigonometric table is not available, the meridian lines can be constructed with the graphic method shown in *Fig. 13-3*:

We take a sheet of blank paper and draw the required number of equally spaced horizontal lines (parallels of latitude). A spacing of 3 - 10 mm per nautical mile is recommended for most applications.

We draw an auxiliary line intersecting the parallels of latitude at an angle numerically equal to the mean latitude. Then we mark the map scale (defined by the spacing of the parallels) periodically on this line, and draw the meridian lines through the points thus located. Compasses can be used to transfer the map scale from the chosen meridian to the auxiliary line.

Small area plotting sheets are available at nautical book stores. Graph Paper Printer (Graphpap) is useful program (shareware) for printing many kinds of graph papers including small-area plotting sheets for any given latitude between 0° and 80°. Copies can be found on the internet, for example at:

http://pharm.kuleuven.be/pharbio/gpaper.htm

Gnomonic Charts

For great circle navigation, the **gnomonic projection** offers the advantage that **any** great circle appears as a straight line. Rhumb lines, however, are curved. A gnomonic chart is obtained by projecting each point on the earth's surface from the earth's center to a plane tangent to the surface. Since the distance of a projected point from the point of tangency varies in proportion with the tangent of the angular distance of the original point from the point of tangency, a gnomonic chart covers less than a hemisphere, and distortions increase rapidly with increasing distance from the point of tangency. In contrast to the Mercator projection, the gnomonic projection is non-conformal (not angle-preserving).

There are three types of gnomonic projection:

If the plane of projection is tangent to the earth at one of the poles (**polar gnomonic chart**), the meridians appear as straight lines radiating from the pole. The parallels of latitude appear as concentric circles. The spacing of the latter increases rapidly as the polar distance increases.

If the point of tangency is on the equator (**equatorial gnomonic chart**), the meridians appear as straight lines parallel to each other. Their spacing increases rapidly as their distance from the point of tangency increases.

The equator appears as a straight line perpendicular to the meridians. All other parallels of latitude (small circles) are lines curved toward the respective pole. Their curvature increases with increasing latitude.

In all other cases (**oblique gnomonic chart**), the meridians appear as straight lines converging at the elevated pole. The equator appears as a straight line perpendicular to the central meridian (the meridian going through the point of tangency). Parallels of latitude are lines curved toward the poles.

Fig. 13-4 shows an example of an oblique gnomonic chart.

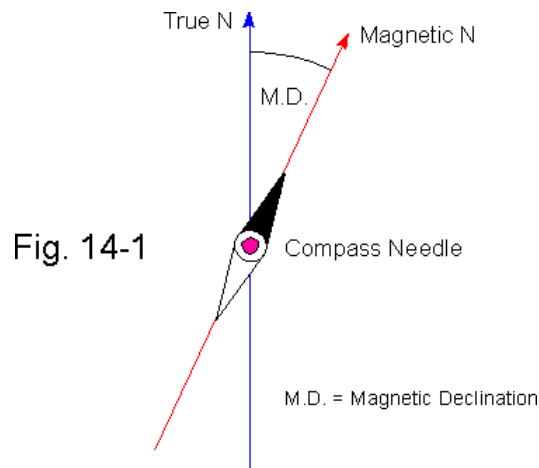
Fig. 13-4

A gnomonic chart is a useful graphic tool for long-distance voyage planning. The intended great circle track is plotted as a straight line from A to B. Obstacles, if existing, become visible at once. The coordinates of the chosen waypoints (preferably those lying on meridian lines) are then read from the graticule and transferred to a Mercator chart, where the waypoints are connected by rhumb line tracks.

Chapter 14

Magnetic Declination

Since the magnetic poles of the earth do not coincide with the geographic poles and due to other irregularities of the earth's magnetic field, the horizontal component of the magnetic field at a given position, called **magnetic meridian**, usually forms an angle with the local geographic meridian. This angle is called **magnetic declination** or, in mariner's language, **magnetic variation**. Accordingly, the needle of a magnetic compass, aligning itself with the local magnetic meridian, does not exactly indicate the direction of true north (*Fig. 14-1*).

Fig. 14-1

Magnetic declination depends on the observer's geographic position and can exceed 30° or even more in some areas. Knowledge of the local magnetic declination is therefore necessary to avoid dangerous navigation errors. Although magnetic declination is often given in the legend of topographic maps, the information may be outdated because magnetic declination varies with time (up to several degrees per decade). In some places, magnetic declination may even differ from official statements due to local distortions of the magnetic field caused by deposits of ferromagnetic ores, etc.

The azimuth formulas described in chapter 4 provide a very useful tool to determine the magnetic declination at a given position. If the observer does not know his exact position, an estimate will suffice in most cases. A sextant is not required for the simple procedure:

1. We choose a celestial body being low in the sky or on the visible horizon, preferably sun or moon. We measure the magnetic compass bearing, B, of the center of the body and note the time. The vicinity of cars, steel objects, magnets, DC power cables, etc. has to be avoided since they distort the magnetic field locally.

2. We extract GHA and Dec of the body from the Nautical Almanac or calculate these quantities with a computer almanac.

3. We calculate the meridian angle, t (or the local hour angle, LHA), from GHA and our longitude (see chapter 4).

4. We calculate the true azimuth, Az_N, of the body from Lat, Dec, and t. The time sight formula (chapter 4) with its accompanying rules is particularly suitable for this purpose since an observed or computed altitude is not needed.

5. Magnetic declination, MD, is obtained by subtracting Az_N from the compass bearing, B.

$$MD = B - Az_N$$

(Add 360° if the angle thus obtained is smaller than –180°. Subtract 360° if the angle is greater than +180°.)

Eastern declination (shown in *Fig. 14-1*) is positive (0°...+180°), western declination negative (0°...–180°).

Chapter 15

Ephemerides of the Sun

The sun is probably the most frequently observed body in celestial navigation. Greenwich hour angle and declination of the sun as well as GHA_{Aries} and EoT can be calculated using the algorithms listed below [8,17]. The formulas are relatively simple and useful for navigational calculations with programmable pocket calculators (10 digits recommended).

First, the time variable, T, has to be calculated from year, month, and day. T is the time, measured in days and dayfractions, before or after Jan 1, 2000, 12:00:00 UT:

$$T = 367 \cdot y - \text{floor}\left\{1.75 \cdot \left[y + \text{floor}\left(\frac{m+9}{12}\right)\right]\right\} + \text{floor}\left(275 \cdot \frac{m}{9}\right) + d + \frac{UT[h]}{24} - 730531.5$$

y is the number of the year (4 digits), m is the number of the month, and d the number of the day in the respective month. UT is Universal Time in decimal format (e.g., 12h 30m 45s = 12.5125). For May 17, 1999, 12:30:45 UT, for example, T is -228.978646. The equation is valid from March 1, 1900 through February 28, 2100.

floor(x) is the greatest integer smaller than x. For example, floor(3.8) = 3, floor(-2.2) = -3. The *floor* function is part of many programming languages, e.g., JavaScript. In general, it is identical with the *int* function used in other languages. However, there seem to be different definitions for the *int* function. This should be checked before programming the above formula.

Mean anomaly of the sun*:

$$g[°] = 0.9856003 \cdot T - 2.472$$

Mean longitude of the sun*:

$$L_M[°] = 0.9856474 \cdot T - 79.53938$$

True longitude of the sun*:

$$L_T[°] = L_M[°] + 1.915 \cdot \sin g + 0.02 \cdot \sin(2 \cdot g)$$

Obliquity of the ecliptic:

$$\varepsilon[°] = 23.439 - 4 \cdot T \cdot 10^{-7}$$

Declination of the sun:

$$Dec[°] = \arcsin(\sin L_T \cdot \sin \varepsilon)$$

Right ascension of the sun (in degrees)*:

$$RA[°] = 2 \cdot \arctan\left(\frac{\cos \varepsilon \cdot \sin L_T}{\cos Dec + \cos L_T}\right)$$

GHA_{Aries}*:

$$GHA_{Aries}[°] = 0.9856474 \cdot T + 15 \cdot UT[h] + 100.46062$$

Greenwich hour angle of the sun*:

$$GHA[°] = GHA_{Aries}[°] - RA[°]$$

*These quantities have to be within the range from 0° to 360°. If necessary, add or subtract 360° or multiples thereof. This can be achieved using the following algorithm which is particularly useful for programmable calculators:

$$y = 360 \cdot \left[\frac{x}{360} - \mathrm{floor}\left(\frac{x}{360} \right) \right]$$

Equation of Time:

$$GAT[h] = \frac{GHA[°]}{15} + 12h$$

(If GAT > 24h, subtract 24h.)

$$EoT[h] = GAT[h] - UT[h]$$

(If EoT > +0.3h, subtract 24h. If EoT < –0.3h, add 24h.)

Semidiameter and Horizontal Parallax

Due to the excentricity of the earth's orbit, semidiameter and horizontal parallax of the sun change periodically during the course of a year. The SD of the sun varies inversely with the distance earth-sun, R:

$$R[AU] = 1.00014 - 0.01671 \cdot \cos g - 0.00014 \cdot \cos(2 \cdot g)$$

(1 AU = $149.6 \cdot 10^6$ km)

$$SD['] = \frac{16.0}{R[AU]}$$

The mean horizontal parallax of the sun is approx. 0.15'. The periodic variation of HP is too small to be of practical significance.

Accuracy

The maximum error of GHA and Dec is about ±0.6'. Results have been cross-checked with *Interactive Computer Ephemeris 0.51* (accurate to approx. 0.1'). Between the years 1900 and 2049, the error was smaller than ±0.3' in most cases (100 dates chosen at random). EoT was accurate to approx. ±2s. In comparison, the maximum error of GHA and Dec extracted from the Nautical Almanac is approx. ±0.25' (for the sun) when using the interpolation tables. The error of SD is smaller than ±0.1'.

Chapter 16

Navigational Errors

Altitude errors

Apart from systematic errors which can be corrected to a large extent (see chapter 2), observed altitudes always contain random errors caused by ,e.g., heavy seas, abnormal atmospheric refraction, and limited optical resolution of the human eye. Although good sextants are manufactured to a mechanical precision of ca. 0.1'- 0.3', the **standard deviation** of an altitude measured with a marine sextant is approximately 1' under fair working conditions. The standard deviation may increase to several arcminutes due to disturbing factors or if a bubble sextant or a plastic sextant is used. Altitudes measured with a theodolite are considerably more accurate (0.1'- 0.2').

Due to the influence of random observation errors, lines of position are more or less indistinct and are better considered as **bands of position**.

Two intersecting bands of position define an **area of position** (ellipse of uncertainty). *Fig. 16-1* illustrates the approximate size and shape of the ellipse of uncertainty for a given pair of position lines. The standard deviations (±x for the first altitude, ±y for the second altitude) are indicated by grey lines.

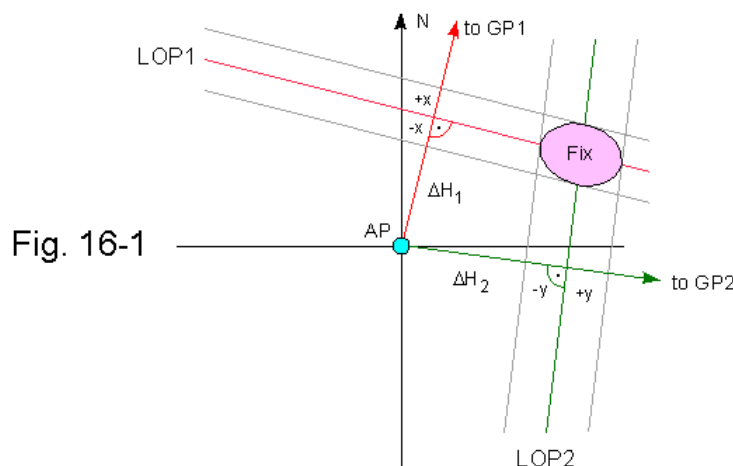



Fig. 16-1

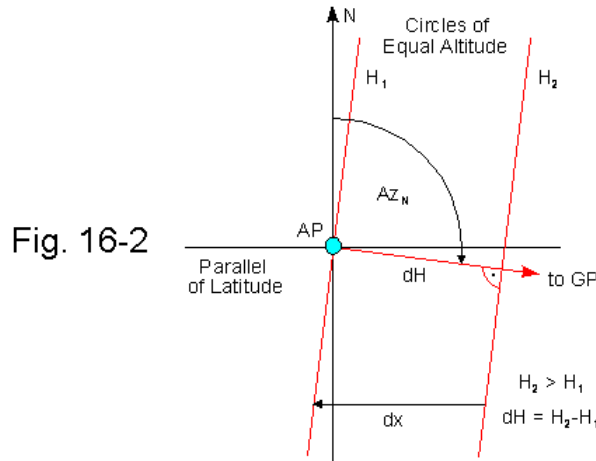
The area of position is smallest if the angle between the bands is 90°. The most probable position is at the center of the area, provided the error distribution is symmetrical. Since position lines are perpendicular to their corresponding azimuth lines, objects should be chosen whose azimuths differ by approx. 90° for best accuracy. An angle between 30° and 150°, however, is tolerable in most cases.

When observing more than two bodies, the azimuths should have a roughly symmetrical distribution (**bearing spread**). With multiple observations, the optimum horizontal angle between two adjacent bodies is obtained by dividing 360° by the number of observed bodies (3 bodies: 120°, 4 bodies: 90°, 5 bodies: 72°, 6 bodies: 60°, etc.).

A symmetrical bearing spread not only improves geometry but also compensates for systematic errors like, e.g., index error.

Moreover, there is an optimum range of altitudes the navigator should choose to obtain reliable results. Low altitudes increase the influence of abnormal refraction (random error), whereas high altitudes, corresponding to circles of equal altitude with small diameters, increase geometric errors due to the curvature of LoP's. The generally recommended range to be used is 20° - 70°, but exceptions are possible.

Time errors

The time error is as important as the altitude error since the navigator usually presets the instrument to a chosen altitude and records the time when the image of the body coincides with the reference line visible in the telescope. The accuracy of time measurement is usually in the range between a fraction of a second and several seconds, depending on the rate of change of altitude and other factors. Time error and altitude error are closely interrelated and can be converted to each other, as shown below (*Fig. 16-2*):

The GP of any celestial body travels westward with an angular velocity of approx. 0.25' per second. This is the rate of change of the LHA of the observed body caused by the earth's rotation. The same applies to each circle of equal altitude surrounding GP (tangents shown in *Fig. 6-2*). The distance between two concentric circles of equal altitude (with the altitudes H_1 and H_2) passing through AP in the time interval dt, measured along the parallel of latitude going through AP is:

$$dx\left[nm\right] = 0.25\cdot\cos Lat_{AP}\cdot dt\left[s\right]$$

dx is also the east-west displacement of a LoP caused by the time error dt. The letter d indicates a small (infinitesimal) change of a quantity (see mathematical literature). $\cos Lat_{AP}$ is the ratio of the circumference of the parallel of latitude going through AP to the circumference of the equator (Lat = 0).

The corresponding difference in altitude (the radial distance between both circles of equal altitude) is:

$$dH\left[{}'\right] = \sin Az_N\cdot dx\left[nm\right]$$

Thus, the rate of change of altitude is:

$$\frac{dH\left[{}'\right]}{dt\left[s\right]} = 0.25\cdot\sin Az_N\cdot\cos Lat_{AP}$$

dH/dt is greatest when the observer is on the equator and decreases to zero as the observer approaches one of the poles. Further, dH/dt is greatest if GP is exactly east of AP (dH/dt positive) or exactly west of AP (dH/dt negative). dH/dt is zero if the azimuth is 0° or 180°. This corresponds to the fact that the altitude of the observed body passes through a minimum or maximum at the instant of meridian transit (dH/dt = 0).

The maximum or minimum of altitude occurs exactly at meridian transit only if the declination of a body is constant. Otherwise, the highest or lowest altitude is observed shortly before or after meridian transit (see chapter 6). The phenomenon is particularly obvious when observing the moon whose declination changes rapidly.

A **chronometer error** is a systematic time error. It influences each line of position in such a way that only the longitude of a fix is affected whereas the latitude remains unchanged, provided the declination does not change significantly (moon!). A chronometer being 1 s fast, for example, displaces a fix by 0.25' to the west, a chronometer being 1 s slow displaces the fix by the same amount to the east. If we know our position, we can calculate the chronometer error from the difference between our true longitude and the longitude found by our observations. If we do not know our longitude, the approximate chronometer error can be found by lunar observations (chapter 7).

Ambiguity

Poor geometry may not only decrease accuracy but may even result in an entirely wrong fix. As the observed horizontal angle (difference in azimuth) between two objects approaches 180°, the distance between the points of intersection of the corresponding circles of equal altitude becomes very small (at exactly 180°, both circles are tangent to each other). Circles of equal altitude with small diameters resulting from high altitudes also contribute to a short distance. A small distance between both points of intersection, however, increases the risk of ambiguity (*Fig. 16-3*).

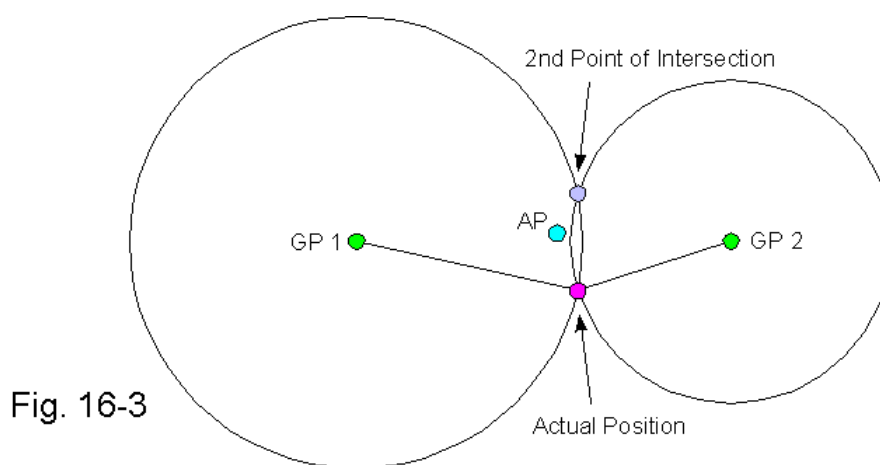



Fig. 16-3

In cases where – due to a horizontal angle near 180° and/or very high altitudes – the distance between both points of intersection is too small, we can not be sure that the assumed position is always close enough to the actual position.

If AP is close to the actual position, the fix obtained by plotting the LoP's (tangents) will be almost identical with the actual position. The accuracy of the fix decreases as the distance of AP from the actual position becomes greater. The distance between fix and actual position increases dramatically as AP approaches the line going through GP1 and GP2 (draw the azimuth lines and tangents mentally). In the worst case, a position error of several hundred or even thousand nm may result !

If AP is exactly on the line going through GP1 and GP2, i.e., equidistant from the actual position and the second point of intersection, the horizontal angle between GP1 and GP2, as viewed from AP, will be 180°. In this case, both LoP's are parallel to each other, and no fix can be found.

As AP approaches the second point of intersection, a fix more or less close to the latter is obtained. Since the actual position and the second point of intersection are symmetrical with respect to the line going through GP1 and GP2, the intercept method can not detect which of both theoretically possible positions is the right one.

Iterative application of the intercept method can only improve the fix if the initial AP is closer to the actual position than to the second point of intersection. Otherwise, an "improved" wrong position will be obtained.

Each navigational scenario should be evaluated critically before deciding if a fix is reliable or not. The distance from AP to the observer's actual position has to be considerably smaller than the distance between actual position and second point of intersection. This is usually the case if the above recommendations regarding altitude, horizontal angle, and distance between AP and actual position are observed.

A simple method to improve the reliability of a fix

Each altitude measured with a sextant, theodolite, or any other device contains systematic and random errors which influence the final result (fix). Systematic errors are more or less eliminated by careful calibration of the instrument used. The influence of random errors decreases if the number of observations is sufficiently large, provided **the error distribution** is symmetrical. Under practical conditions, the number of observations is limited, and the error distribution is more or less unsymmetrical, particularly if an **outlier**, a measurement with an abnormally large error, is present. Therefore, the average result may differ significantly from the true value. When plotting more than two lines of position, the experienced navigator may be able to identify outliers by the shape of the error polygon and remove the associated LoP's. However, the method of least squares, producing an average value, does not recognize outliers and may yield an inaccurate result.

The following simple method takes advantage of the fact that the **median** of a number of measurements is much less influenced by outliers than the **mean** value:

1. We choose a celestial body and measure a series of altitudes. We calculate azimuth and intercept for each observation of said body. The number of measurements in the series has to be odd (3, 5, 7...). The reliability of the method increases with the number of observations.

2. We sort the calculated **intercepts** by magnitude and choose the **median** (the central value in the array of intercepts thus obtained) and its associated azimuth. We discard all other observations of the series.

3. We repeat the above procedure with at least one additional body (or with the same body after its azimuth has become sufficiently different).

4. We plot the lines of position using the azimuth and intercept selected from each series, or use the selected data to calculate the fix with the method of least squares (chapter 4).

The method has been checked with excellent results on land. At sea, where the observer's position usually changes continually, the method has to be modified by advancing AP according to the path of travel between the observations of each series.

Appendix

Literature :

[1] Bowditch, The American Practical Navigator, Pub. No. 9, Defense Mapping Agency Hydrographic/Topographic Center, Bethesda, MD, USA

[2] Jean Meeus, Astronomical Algorithms, Willmann-Bell, Inc., Richmond, VA, USA 1991

[3] Bruce A. Bauer, The Sextant Handbook, International Marine, P.O. Box 220, Camden, ME 04843, USA

[4] Charles H. Cotter, A History of Nautical Astronomy, American Elsevier Publishing Company, Inc., New York, NY, USA (This excellent book is out of print. Used examples may be available at www.amazon.com .)

[5] Charles H. Brown, Nicholl's Concise Guide to the Navigation Examinations, Vol.II, Brown, Son & Ferguson, Ltd., Glasgow, G41 2SG, UK

[6] Helmut Knopp, Astronomische Navigation, Verlag Busse + Seewald GmbH, Herford, Germany (German)

[7] Willi Kahl, Navigation für Expeditionen, Touren, Törns und Reisen, Schettler Travel Publikationen, Hattorf, Germany (German)

[8] Karl-Richard Albrand and Walter Stein, Nautische Tafeln und Formeln (German), DSV-Verlag, Germany

[9] William M. Smart, Textbook on Spherical Astronomy, 6th Edition, Cambridge University Press, 1977

[10] P. K. Seidelman (Editor), Explanatory Supplement to the Astronomical Almanac, University Science Books, Sausalito, CA 94965, USA

[11] Allan E. Bayless, Compact Sight Reduction Table (modified H. O. 211, Ageton's Table), 2nd Edition, Cornell Maritime Press, Centreville, MD 21617, USA

[12] The Nautical Almanac (contains not only ephemeral data but also formulas and tables for sight reduction), US Government Printing Office, Washington, DC 20402, USA

[13] Nautisches Jahrbuch oder Ephemeriden und Tafeln (German), Bundesamt für Seeschiffahrt und Hydrographie, Germany

[14] The Lunar Distance Page, http://www.lunardistance.com

[15] IERS Rapid Service, http://maia.usno.navy.mil

[16] Hannu Karttunen et al., Fundamental Astronomy, 4th Ed., Springer Verlag Berlin Heidelberg New York, 2003

[17] The Astronomical Almanac for the Year 2002, US Government Printing Office, Washington, DC 20402, USA

[18] Michel Vanvaerenberg and Peter Ifland, Line of Position Navigation, Unlimited Publishing, Bloomingdale, Indiana, 2003

[19] George H. Kaplan, The Motion of the Observer in Celestial Navigation, Astronomical Applications Department U. S. Naval Observatory, Washington DC

[20] Ed William's Aviation Formulary, http://williams.best.vwh.net/avform.htm

Author's web site :

http://www.celnav.de